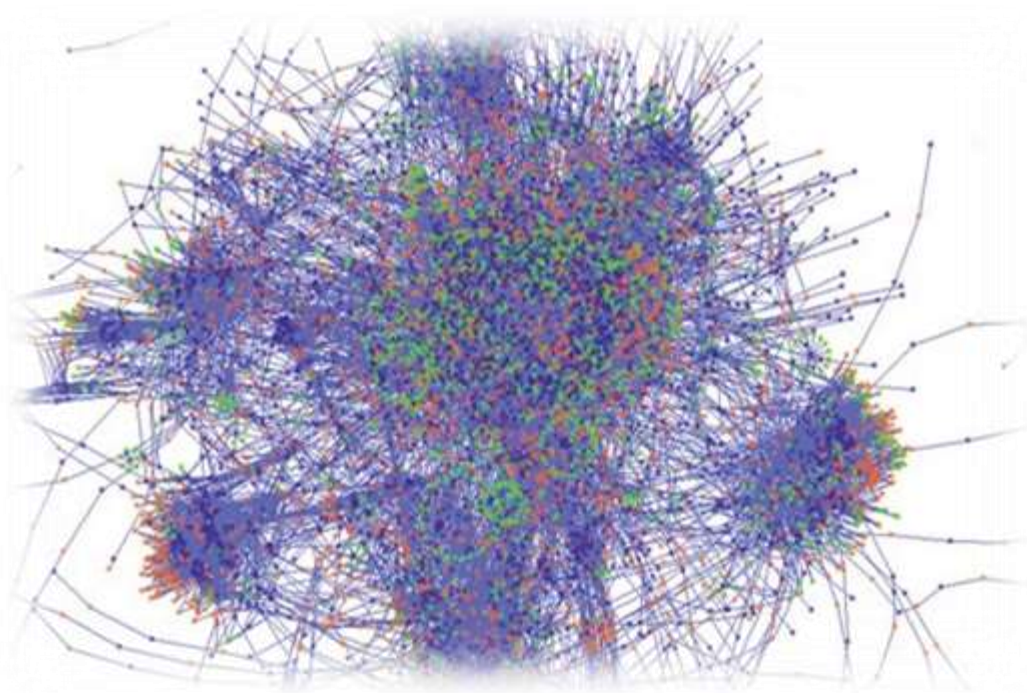




ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΧΗΜΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΙV: ΣΥΝΘΕΣΗΣ & ΑΝΑΠΤΥΞΗΣ ΒΙΟΜΗΧΑΝΙΚΩΝ
ΔΙΑΔΙΚΑΣΙΩΝ

Μεταβολικό μοντέλο γονιδιακής κλίμακας του μικροφύκου *Nannochloropsis gaditana*

Διπλωματική εργασία
Σχινά Αιμιλία



Επιβλέπων καθηγητής: Κολίσης Φραγκίσκος

Ιούνιος 2015

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ	4
ΕΥΧΑΡΙΣΤΙΕΣ	4
ΠΕΡΙΛΗΨΗ	5
ABSTRACT.....	6
ΚΕΦΑΛΑΙΟ 1	7
ΜΕΤΑΒΟΛΙΚΗ ΜΗΧΑΝΙΚΗ,	7
ΣΥΣΤΗΜΙΚΗ ΒΙΟΛΟΓΙΑ.....	7
ΚΑΙ ΜΟΝΤΕΛΑ ΓΟΝΙΔΙΑΚΗΣ ΚΛΙΜΑΚΑΣ	7
1.1 Μεταβολική Μηχανική	7
1.1.1 Εισαγωγή.....	7
1.1.2 Μεταβολική ροή και μεταβολικό μονοπάτι	7
1.1.3 Μεταβολικό δίκτυο.....	10
1.1.4 Εξέλιξη, ζητήματα και διεπιστημονικότητα του κλάδου	13
1.2 Συστημική Βιολογία.....	15
1.3 Μοντέλα γονιδιακής κλίμακας.....	16
1.4 Βιβλιογραφία.....	18
ΚΕΦΑΛΑΙΟ 2	19
ΜΕΤΑΒΟΛΙΚΗ ΜΗΧΑΝΙΚΗ ΒΑΣΙΣΜΕΝΗ ΣΕ ΜΟΝΤΕΛΑ ΓΟΝΙΔΙΑΚΗΣ ΚΛΙΜΑΚΑΣ.....	19
2.1 Μεταβολική Μοντελοποίηση.....	19
2.2 Διαδικασία μοντελοποίησης Γονιδιακής Κλίμακας-Γενικά	21
2.3 Μαθηματική αναπαράσταση μοντέλων γονιδιακής κλίμακας.....	24
2.3.1 Μέθοδοι μαθηματικής αναπαράστασης	25
2.3.2 Βασικά στοιχεία του στοιχειομετρικού πίνακα S.....	26
2.4 Βιβλιογραφία.....	36
ΚΕΦΑΛΑΙΟ 3	37
ΑΝΑΛΥΣΗ ΜΟΝΤΕΛΩΝ ΓΟΝΙΔΙΑΚΗΣ ΚΛΙΜΑΚΑΣ ΥΠΟ ΠΕΡΙΟΡΙΣΜΟΥΣ ..	37
3.1 Μέθοδοι ανάλυσης υπό περιορισμούς	38
3.1.1 Περιορισμοί στις κυτταρικές λειτουργίες.....	38
3.1.2 Μαθηματική αναπαράσταση των περιορισμών.....	39
3.1.3 Εργαλεία ανάλυσης δικτύων	40
3.2 Ανάλυση Ισορροπίας Ροών (Flux Balance Analysis-FBA).....	46
3.3 Βιβλιογραφία.....	50

ΚΕΦΑΛΑΙΟ 4	51
ΤΟ ΜΙΚΡΟΦΥΚΟΣ <i>Nannochloropsis gaditana</i>	51
4.1 Μεταβολισμός φυκών	52
4.2 Καλλιέργεια μικροφυκών	56
4.3 Το θαλάσσιο μικροφύκος <i>Nannochloropsis gaditana</i>	57
4.3.1 Ταξονομία και βασικά χαρακτηριστικά	57
4.3.2 Μεταβολισμός	59
4.3.3 Μεταβολικό προφίλ του <i>Nannochloropsis</i> υπό συνθήκες έλειψης αζώτου	61
4.4 Βιβλιογραφία	63
ΚΕΦΑΛΑΙΟ 5	64
ΚΑΤΑΣΚΕΥΗ ΜΕΤΑΒΟΛΙΚΟΥ ΜΟΝΤΕΛΟΥ ΤΟΥ <i>Nannochloropsis gaditana</i>	64
5.1 Εργαλεία	64
5.2 Διαδικασία κατασκευής μοντέλου	64
5.3 Βιβλιογραφία	114
ΚΕΦΑΛΑΙΟ 6	116
ΕΦΑΡΜΟΓΗ ΑΝΑΛΥΣΗΣ ΡΟΩΝ ΙΣΟΡΡΟΠΙΑΣ ΚΑΙ ΑΝΑΛΥΣΗΣ ΕΥΡΩΣΤΙΑΣ	116
6.1 Αποτελέσματα	116
6.2 Συμπεράσματα	125
6.3 Βιβλιογραφία	127
ΠΑΡΑΡΤΗΜΑ	128
I. Πίνακες με κανόνες γραφής αντιδράσεων και μεταβολιτών	129
II. Χάρτες μεταβολικών μονοπατιών του μοντέλου	132

ΠΡΟΛΟΓΟΣ

Το παρόν τεύχος περιέχει την διπλωματική εργασία της φοιτήτριας της Σχολής Χημικών Μηχανικών ΕΜΠ Σχινά Αιμιλίας. Υπεύθυνος της εργασίας ήταν ο καθηγητής Φραγκίσκος Κολίσης. Μέρος της εργασίας πραγματοποιήθηκε στο Ινστιτούτο Βιολογικών Ερευνών & Βιοτεχνολογίας του Εθνικού Ιδρύματος Ερευνών και συγκεκριμένα στην ερευνητική ομάδα Μεταβολικής Μηχανικής και Βιοπληροφορικής, βοηθώντας του ερευνητή Ελευθέριου Πιλάλη.

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα καταρχήν να ευχαριστήσω όλους όσους συνέβαλαν με οποιονδήποτε τρόπο στην επιτυχή εκπόνηση αυτής της διπλωματικής εργασίας. Θα ήθελα να ευχαριστήσω θερμά τον καθηγητή κ. Φραγκίσκο Κολίση για την ευκαιρία που μου έδωσε να ασχοληθώ με ένα τόσο ενδιαφέρον και απαιτητικό θέμα και για τις πολύτιμες συμβουλές και υποστήριξη του καθ'όλη τη διάρκεια της δουλειάς μου. Θα ήθελα ακόμη να τον ευχαριστήσω για τη δυνατότητα που μου έδωσε να συνεργαστώ με το Εθνικό Ίδρυμα Ερευνών και τον ερευνητή Ελευθέριο Πιλάλη για την πολύτιμη βοήθεια και καθοδήση του καθ'όλη τη διάρκεια εκπόνησης της διπλωματικής εργασίας. Επίσης, ευχαριστώ τον διδακτορικό Μάκη Λαδουκάκη για την πολύτιμη βοήθεια του στον προγραμματισμό με Matlab καθώς και τη διδακτορικό Μαρία Σαββίδου για την αμέριστη βοήθεια της στο θεωρητικό υπόβαθρο της εργασίας και την παροχή δικών της πειραματικών αποτελεσμάτων. Οφείλω σε όλους ένα μεγάλο ευχαριστώ για τον χρόνο που διέθεσαν και για την προθυμία και τις συμβουλές, που ποτέ δε δίστασαν να δώσουν.

ΠΕΡΙΛΗΨΗ

Στην παρούσα διπλωματική εργασία παρουσιάζονται πρώτα οι βασικές αρχές της μεταβολικής μηχανικής και συστημικής βιολογίας και πώς τα μοντέλα γονιδιακής κλίμακας εισάγονται για να ενισχύσουν την *in silico* ανάλυση και το σχεδιασμό σε αυτούς τους τομείς.

Σκοπός είναι η ανάπτυξη ενός μεταβολικού μοντέλου γονιδιακής κλίμακας που να καλύπτει το μεταβολισμό του άλγους *Nannochloropsis gaditana*. Το μοντέλο του *N.gaditana* αποτελεί μία περιεκτική, με βάση τη βιβλιογραφία, γονιδιακής κλίμακας μεταβολική κατασκευή και περιέχει 2205 αντιδράσεις, 2009 μεταβολίτες και 1695 γονίδια. Η κατασκευή διαμερισματοποιήθηκε σε κυτταρόπλασμα, μιτοχόνδριο, χλωροπλάστη, υπεροξειδιδώσωμα και περίπλασμα. Το ενεργό πεδίο του μοντέλου περιλαμβάνει περιλαμβάνει τη γλυκόλυση (κυτταροπλασματική και χλωροπλαστική), τον κύκλο του κιτρικού οξέος, τη φωτοσύνθεση, την οξειδωτική φωσφορυλίωση, τη σύνθεση λιπαρών οξέων, τον κύκλο γλυοξυλικού οξέος, τον κύκλο φωτοαναπνοής, τη βιοσύνθεση χλωροφυλλών και καροτενοειδών, τη σύνθεση πολυαμινών και βιταμινών και αντιδράσεις ζύμωσης. Το μοντέλο του *Nannochloropsis* αξιολογήθηκε μέσα από προσομοιώσεις ανάπτυξης και παραγωγής βιοπροϊόντων, κάτω από μικτές συνθήκες περιβάλλοντος.

Τα αποτελέσματα που προέκυψαν δείχνουν ότι μοντέλο, κάτω από τις κατάλληλες συνθήκες, μπορεί να προσομοιώνει ικανοποιητικά, ποιοτικά όμως, τόσο την ανάπτυξη του κυττάρου του οργανισμού όσο και χαρακτηριστικές μεταβολικές του λειτουργίες, όπως την αύξηση εικοσαπεντανοϊκού σε συνθήκες κορεσμού αζώτου.

ABSTRACT

This thesis addresses basic principles of metabolic engineering and systems biology and the implementation of genome scale metabolic reconstructions in *in silico* analysis and design for systems-metabolic engineering.

The intention of the thesis is the development of a genome scale metabolic network model covering the metabolism of *Nannochloropsis gaditana*. The reconstructed model is a comprehensive literature-based genome scale metabolic reconstruction that accounts for the functions of 2205 reactions, 2009 metabolites and 1695 gene-enzyme-reaction-association entries. The reconstruction was compartmentalized into the cytoplasm, mitochondrion, chloroplast, peroxisome and periplasm. The active scope of the model includes glycolysis, the pentose phosphate pathway (PPP), tricarboxylic acid cycle (TCA cycle), light and dark reactions (Calvin cycle), oxidative phosphorylation, fatty acid synthesis, beta-oxidation, glyoxylate cycle, photorespiratory cycle, chlorophylls and carotenoid synthesis, biosynthesis of polyamines and vitamins and fermentative reactions. The model was validated through the simulation of growth and bioproducts production under mixotrophic conditions.

The results of evaluation process show that the model can predict satisfactorily, in a qualitative way the growth of the algal cell as well as distinct metabolic functions, such as the increased production of EPA in nitrogen-repletion conditions.

ΚΕΦΑΛΑΙΟ 1

ΜΕΤΑΒΟΛΙΚΗ ΜΗΧΑΝΙΚΗ, ΣΥΣΤΗΜΙΚΗ ΒΙΟΛΟΓΙΑ ΚΑΙ ΜΟΝΤΕΛΑ ΓΟΝΙΔΙΑΚΗΣ ΚΛΙΜΑΚΑΣ

1.1 Μεταβολική Μηχανική

1.1.1 Εισαγωγή

Η μεταβολική μηχανική αποτελεί ένα σύγχρονο διεπιστημονικό κλάδο της μηχανικής ο οποίος ασχολείται με την ανάλυση και τροποποίηση μεταβολικών μονοπατιών και κατ' επέκταση δικτύων και έχει ραγδαία εξέλιξη τα τελευταία είκοσι χρόνια. Μέσω των τεχνικών της μοριακής βιολογίας και της μηχανικής χημικών διεργασιών η μεταβολική μηχανική αποτελεί πλέον κομβικό σημείο ερευνητικής δραστηριότητας στη χημική μηχανική, βιοχημική μηχανική, εφαρμοσμένη μικροβιολογία και κυτταρική φυσιολογία.

Οι μέθοδοι χειραγώγησης μεταβολικών μονοπατιών βασίζονται κυρίως στη χρήση χημικών μεταλλαξιογόνων και τεχνικών διαλογής των ισχυρότερων στελεχών για την επίτευξη συγκεκριμένων σκοπών. Παρά την επιτυχία και την ευρεία αποδοχή τους, τα γενετικά και μεταβολικά προφίλ των μεταλλαγμένων στελεχών ήταν πολύ πτωχά χαρακτηρισμένα και οι μεταλλαξιογενέσεις αντιμετωπίζονταν ως τυχαίες διαδικασίες χωρίς σημαντικό επιστημονικό υπόβαθρο.

Σημαντική ώθηση στη χειραγώγηση των μονοπατιών έδωσε η ανάπτυξη τεχνικών της μοριακής βιολογίας για τον ανασυνδυασμό του DNA. Η γενετική μηχανική επέτρεψε την τροποποίηση συγκεκριμένων ενζυμικών αντιδράσεων στα μεταβολικά μονοπάτια, με αποτέλεσμα την κατασκευή καλά ορισμένων γενετικών υποβάθρων. Η μεταβολική μηχανική μπορεί να οριστεί ως: *η στοχευμένη βελτίωση της παραγωγής προϊόντων ή κυτταρικών ιδιοτήτων μέσω της τροποποίησης συγκεκριμένων βιοχημικών αντιδράσεων ή μέσω της εισαγωγής νέων αντιδράσεων χρησιμοποιώντας την τεχνολογία του ανασυνδυασμένου DNA.*

1.1.2 Μεταβολική ροή και μεταβολικό μονοπάτι

Μία καινοτόμος πλευρά της μεταβολικής μηχανικής είναι σημασία που δίνεται στα ολοκληρωμένα μεταβολικά μονοπάτια αντί στις μεμονωμένες αντιδράσεις. Έτσι, μελετούνται ολόκληρα δίκτυα βιοχημικών αντιδράσεων και ζητήματα σύνθεσης

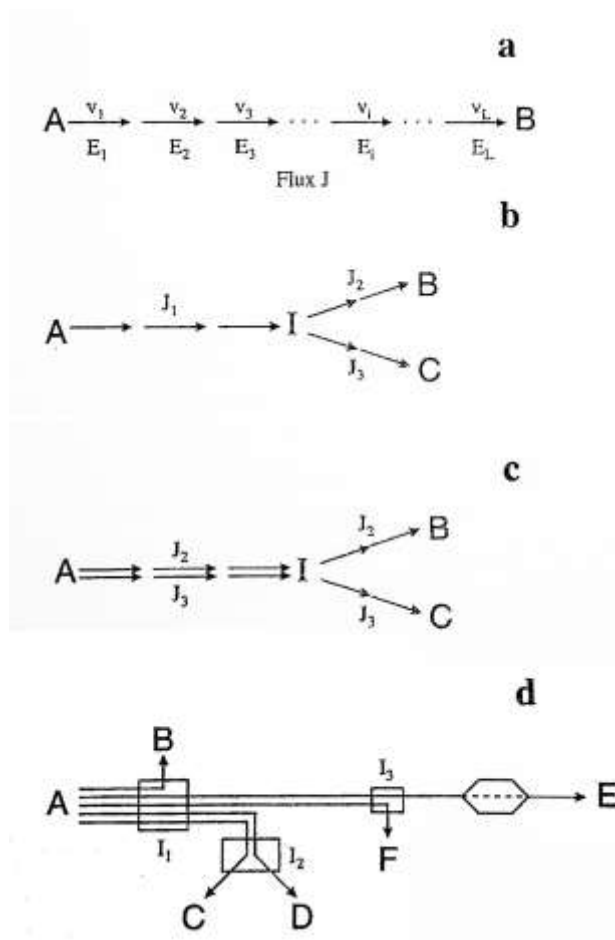
μονοπατιών, θερμοδυναμικής καθώς και ελέγχου της ροής μονοπατιού. Μία από τις σημαντικότερες έννοιες στη μεταβολική μηχανική είναι η μεταβολική ροή. Η έννοια αυτή και η ρύθμισή της μελετάται εδώ και σαράντα χρόνια από αρκετούς βιοχημικούς και τα πορίσματα της έρευνας βρίσκουν εφαρμογή στην μεταβολική μηχανική. Μπορούμε να πούμε ότι ο συνδυασμός αναλυτικών τεχνικών για την ποσοτικοποίηση των ροών και της ρύθμισής τους, μαζί με τεχνικές μοριακής βιολογίας για την εφαρμογή προτεινόμενων γενετικών τροποποιήσεων, αποτελούν την ουσία της μεταβολικής μηχανικής.

Όταν τα ανωτέρω χρησιμοποιηθούν με επαναληπτικό τρόπο, προκύπτει μία ισχυρή μέθοδος για την συστηματική βελτίωση κυτταρικών ιδιοτήτων σε ένα ευρύ φάσμα εφαρμογών.

Η ροή αποτελεί ένα θεμελιώδη παράγοντα καθορισμού της φυσιολογίας του κυττάρου και την πιο κρίσιμη παράμετρο ενός μεταβολικού μονοπατιού. Καθώς τα μεταβολικά μονοπάτια (ή οδοί) και οι ροές τους βρίσκονται στο επίκεντρο της μεταβολικής μηχανικής είναι σημαντικό να δοθούν οι ορισμοί τους και να εξηγηθεί κατά το δυνατόν η σημασία τους. Ως μεταβολικό μονοπάτι ορίζεται μία ακολουθία πραγματοποιήσιμων και παρατηρήσιμων βιοχημικών, ενζυμικά καταλυόμενων αντιδράσεων που ενώνουν ένα σύνολο μεταβολιτών εισόδου και εξόδου. Συζευγμένα μεταβολικά μονοπάτια απαρτίζουν ένα μεταβολικό δίκτυο. Ο ρυθμός με τον οποίο οι μεταβολίτες εισόδου παράγουν μεταβολίτες εξόδου αποτελεί τη μεταβολική ροή του μονοπατιού.

Για το γραμμικό μονοπάτι της Εικ. 1a, η ροή του, J , ισούται με τους ρυθμούς των επιμέρους αντιδράσεων σε *σταθερή κατάσταση*. Προφανώς, η σταθερή κατάσταση επιτυγχάνεται όταν μετά από αυξομείωση των συγκεντρώσεων των ενδιάμεσων μεταβολιτών προκύψει εξίσωση των ρυθμών των αντιδράσεων ($v_1 = v_2 = \dots v_i \dots = v_L$). Κατά τη διάρκεια μιας μεταβατικής κατάστασης, οι επιμέρους ρυθμοί δεν είναι ίσοι μεταξύ τους και η ροή του μονοπατιού είναι μεταβλητή και όχι σαφώς καθορισμένη. Για το διακλαδισμένο μονοπάτι της Εικ. 1b (διάσπαση στο ενδιάμεσο I) εμφανίζονται δύο επιπλέον ροές για κάθε ένα από τα δύο μονοπάτια-κλάδους. Σε σταθερή κατάσταση οι ροές συνδέονται μέσω της σχέσης: $J_1 = J_2 + J_3$, με τις ροές κάθε κλάδου να ισούνται με τους αντίστοιχους ρυθμούς αντιδράσεων του κλάδου. Είναι συχνά βολικό η ροή J_1 να θεωρείται ως ο συνδυασμός των ροών J_2 και J_3 (Εικ.

1c). Με τον τρόπο αυτό, ένα πολύπλοκο δίκτυο (Εικ. 1d) μπορεί να διασπαστεί σε αρκετά γραμμικά μονοπάτια, καθένα από τα οποία έχει τη δική του ροή. Πρέπει να σημειωθεί ότι για κάθε μονοπάτι της Εικ. 1 μία αναγκαία συνθήκη για την επίτευξη σταθερής κατάστασης είναι οι ρυθμοί των αρχικών και τελικών αντιδράσεων (ή ισοδύναμα, οι συγκεντρώσεις των αρχικών και τελικών μεταβολιτών A, B και C αντίστοιχα) να είναι σταθεροί. Αυτό συνήθως επιτυγχάνεται διατηρώντας σταθερές τις εξωκυτταρικές συγκεντρώσεις των μεταβολιτών σε ένα βιοαντιδραστήρα συνεχούς λειτουργίας (συνεχής καλλιέργεια – chemostat).



Εικόνα 1.1 Παραδείγματα απλών μονοπατιών

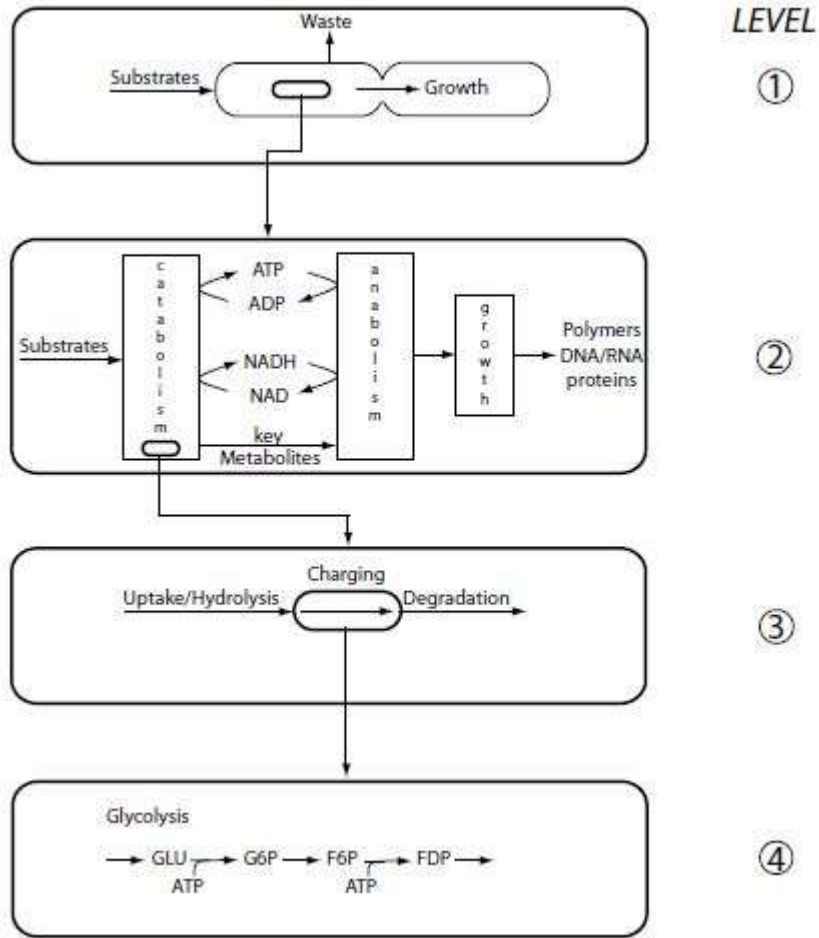
Τα σημαντικά σημεία στους ορισμούς μεταβολικού μονοπατιού και μεταβολικής ροής είναι η *εφικτότητα* και η *παρατηρησιμότητα*. Αφενός θα πρέπει να αντιστοιχούν σε πραγματικές, ενζυμικά καταλυόμενες αντιδράσεις και αφετέρου οι ροές τους θα πρέπει να μπορούν να παρατηρηθούν και να μετρηθούν πειραματικά. Πολύ συχνά

υπάρχουν περισσότερες από μία ακολουθίες αντιδράσεων μεταξύ συγκεκριμένων μεταβολιτών εισόδου και εξόδου. Αν όμως οι ροές των ενδιάμεσων αυτών ακολουθιών δεν μπορούν ανεξάρτητα να προσδιοριστούν, τότε η ένταξή τους στο δίκτυο δεν θα προσφέρει επιπλέον χρήσιμες πληροφορίες. Για τον λόγο αυτό είναι συνήθως προτιμότερο οι αντιδράσεις αυτές να συμπυκνωθούν σε μικρότερο αριθμό μονοπατιών των οποίων οι ροές μπορούν να παρατηρηθούν. Αντικείμενο συνεχούς έρευνας είναι η προσπάθεια (π.χ. μέσω πειραματικών μετρήσεων) να διαφωτιστεί η λειτουργία των συνεπτυγμένων αυτών μονοπατιών καθώς με τον τρόπο αυτό αυξάνεται η ανάλυση (resolution) του βιοχημικού δικτύου. Η έννοια της μεταβολικής ροής, των μεθόδων για τον προσδιορισμό των ροών *in vivo*, η υπολογιστική μοντελοποίηση των μεταβολικών δικτύων *in silico*, και των συμπερασμάτων που μπορούν να εξαχθούν από τη συστηματική μελέτη των μεταβολικών ροών αποκαλούνται συλλήβδην *ανάλυση μεταβολικών ροών (Metabolic Flux Analysis – MFA)*. (Stephanopoulos, et al., 1998)

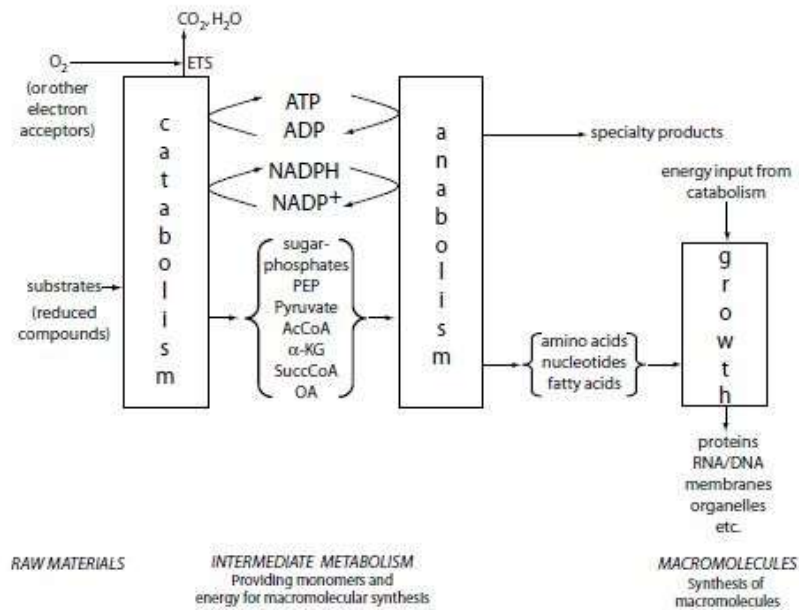
1.1.3 Μεταβολικό δίκτυο

Η λειτουργία των κυττάρων βασίζεται σε πολύπλοκα δίκτυα αλληλεπιδρούντων χημικών αντιδράσεων οι οποίες είναι προσεκτικά οργανωμένες στον χώρο και το χρόνο. Αυτά τα δίκτυα βιοχημικών αντιδράσεων παράγουν παρατηρήσιμες κυτταρικές λειτουργίες.

Τα μοντέλα γονιδιακή κλίμακας των μεταβολικών δικτύων περιλαμβάνουν εκατοντάδες μεταβολίτες και πολλές φορές πάνω από χίλιες αντιδράσεις. Η λειτουργία αυτών των δικτύων είναι δυσνόητη για τον άνθρωπο. Για αυτό το λόγο γίνεται ιεράρχηση των ιδιοτήτων των δικτύων προς απλοποίηση της αντίληψης των λειτουργιών αυτών. Στην Εικ. 2 απεικονίζεται η ιεραρχική αποδόμηση των λειτουργιών των μεταβολικών δικτύων. Στην Εικ. 3 απεικονίζονται τα βασικά τμήματα του κυτταρικού μεταβολισμού.



Εικόνα 1.2 Λειτουργική αποδόμηση του μεταβολισμού σε τέσσερα επίπεδα. Επίπεδο 1: ολόκληρο κύτταρο; Επίπεδο 2: μεταβολικοί τομείς; Επίπεδο 3: μονοπάτια; Επίπεδο 4 : μεμονωμένες αντιδράσεις.



Εικόνα 1.3 Βασικά τμήματα του κυτταρικού μεταβολισμού.

Επίπεδο 1: Είσοδοι και έξοδοι κυττάρου. Συνολικά, ο μεταβολισμός περιλαμβάνει τις ενζυμικές αντιδράσεις που αφορούν τη μετατροπή μορίων-υποστρωμάτων στα απαραίτητα δομικά στοιχεία των μακρομορίων και άλλων ζωτικών προϊόντων για την ανάπτυξη και συντήρηση. Έτσι μία πιο «αδρή» περιγραφή της συνολικής δραστηριότητας του μεταβολισμού περιλαμβάνει τα υποστρώματα ως εισόδους και τη βιομάζα και τα μεταβολικά παραπροϊόντα ως εξόδους.

Επίπεδο 2: Τομείς. Μία πιο κομψή ματιά στον μεταβολισμό αποκαλύπτει ότι μπορεί να διαχωριστεί σε δύο βασικού τομείς, όπως φαίνεται στην Εικ. 3. Ο καταβολισμός εκτελεί την αποδόμηση των υποστρωμάτων μέσω μιας σειράς συγκλίνοντων μονοπατιών που οδηγούν σε έντεκα μεταβολίτες κεντρικής σημασίας, οι οποίοι ονομάζονται βιοσυνθετικοί πρόδρομοι. Ο αναβολισμός είναι ένα σύνολο αποκλίνοντων μονοπατιών που προέρχονται από αυτούς τους κεντρικούς μεταβολίτες προς σχηματισμό μονομερών δομικών στοιχείων για τη μακρομοριακή βιοσύνθεση.

Επίπεδο 3: Μονοπάτια. Μια ακόμα πιο υψηλή ανάλυση αποκαλύπτει ότι τα μονοπάτια εξυπηρετούν ένα σημαντικό ρόλο. Για παράδειγμα, ο καταβολισμός των κύριων κατηγοριών βιομορίων ακολουθεί το ίδιο μοτίβο. Πρώτα, τα υποστρώματα συλλέγονται από το κύτταρο, υδρολύονται εφόσον χρειάζεται., ενεργοποιούνται από ένα συμπάραγοντα και ύστερα αποδομούνται για να αποδώσουν ενέργεια και άλλες ιδιότητες αποθηκευμένες στα μόρια-μεταφορείς. Σε αυτό το επίπεδο περιγραφής, τα βασικά στοιχεία του μεταβολισμού αρχίζουν να εξαρτούνται σε βασικές χημικές αρχές όπως τη στοιχειομετρική δομή και την κινητική ρύθμιση.

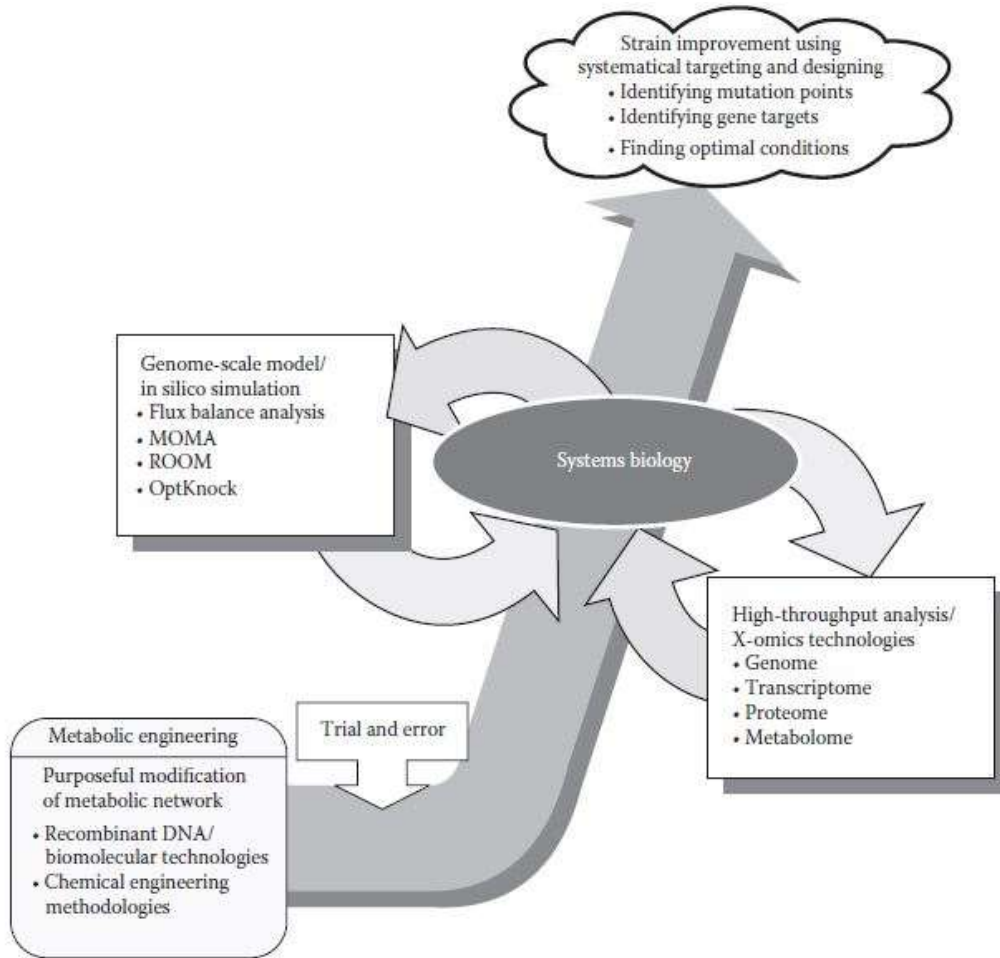
Επίπεδο 4: Μεμονωμένες αντιδράσεις. Στο πιο αναλυτικό στάδιο περιγραφής μελετούνται όλοι οι βιοχημικοί μετασχηματισμοί που πραγματοποιούνται στο κύτταρο. Υψηλής απόδοσης δεδομένα παρέχουν τις απαραίτητες πληροφορίες για την περιγραφή των κυττάρων σε αυτό το στάδιο. Οι βιοχημικοί μετασχηματισμοί διαίρονται σε λίγες βασικές κατηγορίες. Κάποια παραδείγματα αυτών περιλαμβάνουν την τρανσαμίνωση, φωσφορυλίωση, ισομερίωση, αφυδάτωση και αυτοξειδοαναγωγή. Έτσι, υπάρχουν χημικοί «κανόνες» οι οποίοι υπαγορεύουν το είδος συνδέσεων που υπάρχει στα μεταβολικά δίκτυα. Οι βιοχημικοί έχουν επινοήσει την ονοματολογία που κατηγοριοποιεί τις μετατροπές και έναν αριθμό, Enzyme Commission number(E.C.), ο οποίος συνδέεται με κάθε ενζυμικά καταλυόμενη

αντίδραση. Επιπλέον υπάρχουν θερμοδυναμικοί περιορισμοί που σχετίζονται με αυτούς του μετασχηματισμούς και υποδεικνύουν τη θερμοδυναμική εφικτότητα μίας αντίδρασης και την κατάσταση ισορροπίας. (Palsson, 2006)

1.1.4 Εξέλιξη, ζητήματα και διεπιστημονικότητα του κλάδου

Ένα από τα βασικά χαρακτηριστικά της μεταβολικής μηχανικής είναι η τροποποίηση των μεταβολικών δικτύων μέσω προσθήκης, διαγραφής ή/και μετατροπής των μεταβολικών μονοπατιών με στόχο τη βελτιωμένη παραγωγή συγκεκριμένων μεταβολιτών. Ωστόσο, λόγω της πολυσύνθετης φύσης των μεταβολικών δικτύων και της ρύθμισής τους, οι χειρισμοί των μεταβολικών μονοπατιών μπορεί να έχουν απρόβλεπτες συνέπειες και να παρεμποδίζουν την επίτευξη του αρχικού στόχου. Προγενέστερα της έλευσης των συστημικών προσεγγίσεων, η μεταβολική μηχανική βασιζόταν στη διαίσθηση και τη βιοχημική γνώση για την επιλογή μεταβολικών μονοπατιών προς τροποποίηση. Τα αποτελέσματα αυτής της προσέγγισης ήταν συχνά απρόοπτα και τα στελέχη που προέκυπταν απαιτούσαν εκτενείς μικρορυθμίσεις για να αποφέρουν βιώσιμα, παραγωγικά στελέχη. Η ενσωμάτωση πολύπλοκων μεθόδων μεταβολικής μηχανικής περιλαμβάνει γενετικές τροποποιήσεις που σχετίζονται με σημαντικές φαινοτυπικές αλλαγές στον οργανισμό. Τέτοιες αλλαγές έχουν ως αποτέλεσμα χαμηλότερους ρυθμούς ανάπτυξης και παραγωγής περιττών και πιθανώς τοξικών παραπροϊόντων. Λόγω αυτών των προβλημάτων, οι κλασικές προσεγγίσεις της μεταβολικής μηχανικής είναι συχνά χρονοβόρες, απαιτούν μεγάλο φόρτο εργασίας και είναι αναποτελεσματικές από οικονομική σκοπιά.

Τα τελευταία χρόνια, η μεταβολική μηχανική έχει στραφεί στην ενσωμάτωση συστημικών προσεγγίσεων, οι οποίες βασίζονται κυρίως σε μεγάλης κλίμακας πειράματα καθώς και σε υπολογιστική ανάλυση μεταβολικών και ρυθμιστικών δικτύων. Υπάρχει πλέον έντονο ενδιαφέρον μεταξύ επιστημόνων και μηχανικών για τη μελέτη των κυττάρων και των μικροοργανισμών στο πλαίσιο της συστημικής βιολογίας.



Εικόνα 1.4 Μεταβολική μηχανική βασισμένη σε συστημικές προσεγγίσεις, υψηλής απόδοσης τεχνικές και μοντέλα γονιδιακής κλίμακας.

Οι τεχνολογίες υψηλής απόδοσης μαζί με τις τεχνικές αλληλουχίας γονιδιωμάτων επιτρέπουν την κατασκευή εκτενών βάσεων δεδομένων που αντιπροσωπεύουν τη γνώση των μεταβολικών δικτύων συγκεκριμένων οργανισμών, όπως η βάση KEGG (Kyoto Encyclopedia of Genes and Genomes), SEED. Αυτές οι βάσεις περιέχουν βιοχημικές, μοριακές και γενετικές πληροφορίες και συμβάλλουν στην εξέλιξη πιο συστηματικής και αποτελεσματικής μεταβολικής μηχανικής. Η κατασκευή *in silico* στοιχειομετρικών μεταβολικών μοντέλων γονιδιακής κλίμακας εμφανίστηκε χάρη στην εισροή δεδομένων υψηλής απόδοσης.

1.2 Συστημική Βιολογία

Παραδοσιακά οι βιολόγοι χρησιμοποιούσαν για την ανάλυση συστημάτων την προσέγγιση bottom-up: το σύστημα αποδομείται και αναλύεται στα στοιχειώδη συστατικά του και βάσει των ιδιοτήτων τους αξιολογείται η συστημική συμπεριφορά. Παρόλο που η κατανόηση της συμπεριφοράς και του ρόλου κάθε μεμονωμένου συστατικού σε ένα συγκεκριμένο βιολογικό σύστημα είναι σημαντική, η απομονωμένη μελέτη κάθε συστατικού δεν μπορεί να αποδώσει την ολική εικόνα λειτουργίας του συστήματος. Για μια πιο ολοκληρωμένη γνώση της συμπεριφοράς του συστήματος, θα πρέπει να μελετηθούν ταυτόχρονα διάφορα από τα συστατικά του. Η συστημική βιολογία αναζητά την ενοποίηση της υπάρχουσας γνώσης βιολογίας ενός συγκεκριμένου συστήματος με ποσοτικά πειράματα υψηλής απόδοσης για να δια φωτίσει πώς διαφορετικά υποσυστήματα αλληλοεπιδρούν και λειτουργούν ως σύνολο.

Με την πρόοδο των τεχνικών υψηλής απόδοσης, οι πληροφορίες για τα μοριακά χαρακτηριστικά των κυττάρων αυξάνονται ραγδαία. Αυτό δημιουργεί την ανάγκη ανάπτυξης μεθόδων εξαγωγής πολύτιμων πληροφοριών από μεγάλης κλίμακας σύνολα δεδομένων. Επιπλέον, πρέπει να αναπτυχθούν μέθοδοι που συνδέουν αυτές τις πληροφορίες με κυτταρικούς φαινότυπους. Για παράδειγμα, σε γονιδιακά δεδομένα μπορεί να αναγνωρίζονται συγκεκριμένα αλληλόμορφα που αυξάνουν την ευαισθησία σε ορισμένες ασθένειες, αλλά δεν αποκαλύπτουν τους βιολογικούς μηχανισμούς που προκαλούν την αύξηση της ευαισθησίας. Μόνο με συνδυασμό των γονιδιακών πληροφοριών με τη γνώση των δομών μεταβολισμού, ρύθμισης και σηματοδότησης του δικτύου υπάρχει δυνατότητα προσδιορισμού του τρόπου με τον οποίο συγκεκριμένες γενετικές μεταβολές προκαλούν τα παρατηρούμενα φαινοτυπικά αποτελέσματα (π.χ. ευαισθησία σε ασθένειες).

Μεγάλος όγκος πληροφοριών που χαρακτηρίζουν μικροοργανισμούς οι οποίοι χρησιμοποιούνται συχνά σε εφαρμογές μεταβολικής μηχανικής είναι πλέον διαθέσιμος σε μορφή μεταγραφικών, πρωτεόμικών, μεταβολομικών και φαινοτυπικών δεδομένων. Προκειμένου να μεγιστοποιηθεί η αξία των πληροφοριών που εξάγεται, θα πρέπει όλα αυτά τα σύνολα δεδομένων να αναλυθούν μαζί, ενώ κάθε ένα μόνο του επιτρέπει τη μελέτη μια συγκεκριμένης όψης της συνολικής μικροβιακής φυσιολογίας. Η συστημική βιολογία αναζητά την επίτευξη αυτού του

στόχου μέσω της κατασκευής περιεκτικών μοντέλων των βιολογικών δικτύων. Αυτά τα μοντέλα μπορούν να χρησιμοποιηθούν ως πλαίσια ενσωμάτωσης δεδομένων προς διευκόλυνση επιστημονικών ανακαλύψεων και υποθέσεων. Υπάρχει δυνατότητα αναπαράστασης μηχανισμών όπως εναλλακτικά μεταβολικά μονοπάτια, αποτελέσματα ανατροφοδότησης στη μεταγραφική ρύθμιση και αλληλεπιδράσεις σηματοδότησης. Η ικανότητα αυτή, του να συμπεριλαμβάνονται συστηματικά αυτοί οι πολύπλοκοι μηχανισμοί μπορεί να βελτιώσει σημαντικά την ικανότητα εκμετάλλευσης και σχεδιασμού βακτηρίων για την παραγωγή επιθυμητών βιοπροϊόντων.

1.3 Μοντέλα γονιδιακής κλίμακας

Η διαδικασία κατασκευής μοντέλων γονιδιακής κλίμακας είναι ένα ουσιαστικό πρώτο βήμα προς την ανάλυση της μεταβολικής φυσιολογίας μικροβιακών οργανισμών. Τα μεταβολικά μοντέλα επιτρέπουν τον τοπολογικό χαρακτηρισμό του δικτύου, τον προσδιορισμό σημαντικών γονιδίων, στόχους διαγραφής γονιδίων για βελτιωμένη παραγωγή παραπροϊόντων καθώς και πρόβλεψη φαινότυπων ανάπτυξης κάτω από ποικίλες συνθήκες. Επιπλέον, η *in silico* μοντελοποίηση διευκολύνει την ανάλυση διάφορου τύπου συνόλου δεδομένων, όπως χαρακτηριστικά στοιχεία έκφρασης γονιδίων και πρωτεϊνών αλλά και την απεικόνιση αυτών των δεδομένων μέσα σε ένα λειτουργικό πλαίσιο του μοντέλου.

Μία από τις πιο ελπιδοφόρες πτυχές των *in silico* μεταβολικών μοντέλων είναι η ικανότητα τους να προβλέπουν με ακρίβεια το φαινότυπο του οργανισμού βάσει μόνο του μεταβολικού δικτύου και χωρίς να βασίζονται στη γνώση των μηχανισμών ρύθμισης. Επιτυχείς προσπάθειες κατευθύνονται προς τη συστηματική επέκταση των μεταβολικών δικτύων και προς την ανάπτυξη νέων μεθόδων για τη βελτίωση των *in silico* φαινοτυπικών προβλέψεων από τα μοντέλα γονιδιακής κλίμακας.

Η ραγδαία ανάπτυξη κατασκευής μεταβολικών δικτύων και μεθόδων *in silico* προσομοιώσεων έχει οδηγήσει στην αναγνώριση της αξίας των μοντέλων γονιδιακής κλίμακας για τη μεταβολική μηχανική. Οι σύνθετες αλληλεπιδράσεις μεταξύ γονιδίων ακόμα και σε απλα μικροβιακά κύτταρα περιπλέκουν την ταυτοποίηση των σωστών γονιδίων προς χειρισμό για να επιτευχθεί ο επιθυμητός φαινότυπος. Ωστόσο, τα μοντέλα γονιδιακής κλίμακας συμβάλλουν στο να γίνει η διαδικασία ταυτοποίησης

σημαντικά πιο γρήγορα και εύκολα και έτσι επιτρέπουν τη γρήγορη ιεράρχηση γενετικών χειρισμών. Παρά το γεγονός ότι η σχέση γονότυπου-φαινότυπου είναι αρκετά σύνθετη και μη γραμμική και περιπλέκεται ακόμα περισσότερο από τους μηχανισμούς ρύθμισης, έχουν αναφερθεί αρκετές επιτυχείς μελέτες που στοχεύουν στην αναγνώριση γενετικών χειρισμών βασισόμενες σε μοντέλα γονιδιακής κλίμακας. (Lee, et al., 2010)

1.4 Βιβλιογραφία

Lee, S. Y. et al., 2010. Systems Biology, Genome-Scale Models and Metabolic Engineering. In: C. D. Smolke, ed. *The metabolic pathway engineering handbook: Tools and applications*. Boca Raton: CRC Press, pp. 338-341.

Palsson, B. Ø., 2006. Metabolic Networks. In: *Systems Biology (Properties of Reconstructed Networks)*. New York: Cambridge University Press, pp. 29-39.

Stephanopoulos, G. N., Aristidou, A. A. & Nielsen, J., 1998. The Essence of Metabolic Engineering. In: *Metabolic Engineering (Principles and methodologies)*. San Diego: Academic Press, pp. 9-20.

ΚΕΦΑΛΑΙΟ 2

ΜΕΤΑΒΟΛΙΚΗ ΜΗΧΑΝΙΚΗ ΒΑΣΙΣΜΕΝΗ ΣΕ ΜΟΝΤΕΛΑ ΓΟΝΙΔΙΑΚΗΣ ΚΛΙΜΑΚΑΣ

Μία από τις μεγαλύτερες προκλήσεις στην προσπάθεια υπερπαραγωγής επιθυμητών υποπροϊόντων είναι η αναγνώριση γονιδίων που πρέπει να χειραγωγηθούν για να επιτευχθεί ο επιθυμητός φαινότυπος. Η επιλογή της πιο παραγωγικής στρατηγικής γενετικής χειραγώγησης απαιτεί κατανόηση του πώς το τροποποιημένο μονοπάτι θα λειτουργήσει στο πλαίσιο του συνολικού συστήματος. Οι προσεγγίσεις δεδομένων υψηλής απόδοσης τυπικά επιτρέπουν την αναγνώριση δυνητικών σημείων σε ήδη υπάρχοντα μηχανικά τροποποιημένα στελέχη, αλλά είναι λιγότερο χρήσιμες στην επιλογή γονιδίων προς υπερέκφραση ή διαγραφή. Σε αυτή τη σχεδιαστική φάση της μεταβολική μηχανικής τα μοντέλα γονιδιακής κλίμακας είναι πολλά υποσχόμενα. Οι τεχνικές μοντελοποίησης υπό περιορισμούς, όπως η ανάλυση ισορροπίας ροών (Flux Balance Analysis-FBA), μπορούν να εφαρμοστούν σε μεταβολικά μοντέλα γονιδιακής κλίμακας και προσφέρουν καλύτερη ενόραση στην αλληλεπίδραση των μεταβολικών μονοπατιών μέσα στο *in vivo* σύστημα. Καθώς και γρήγορη αξιολόγηση δυνητικών μηχανικών στρατηγικών. Η όλο και αυξανόμενη διαθεσιμότητα των υψηλής ποιότητας μοντέλων γονιδιακής κλίμακας μαζί με την ανάπτυξη νέων μεθόδων μοντελοποίησης υπό περιορισμούς έχει συμβάλλει στην αυξανόμενη εφαρμογή τέτοιων μοντέλων σε μηχανικά μεταλλαγμένους μικροοργανισμούς για να υπερπαραγάγουν εμπορικά επιθυμητά μεταβολικά προϊόντα. (Lee, et al., 2010)

2.1 Μεταβολική Μοντελοποίηση

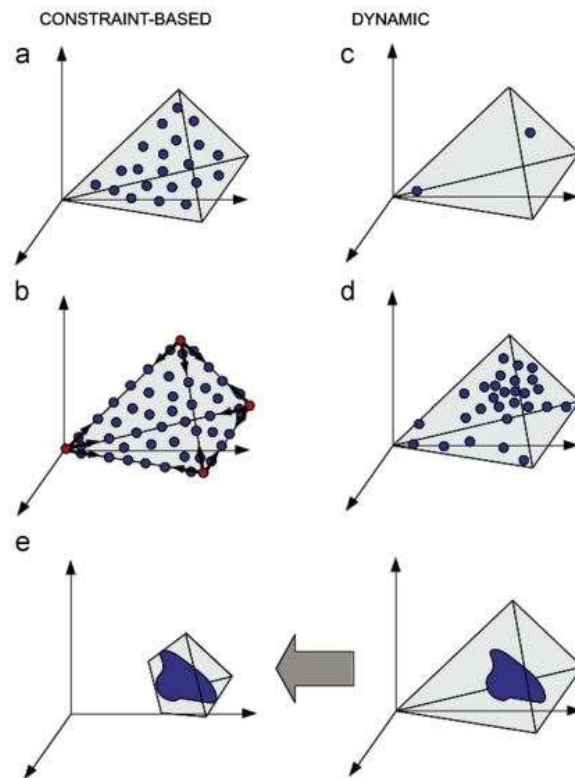
Η μεταβολική μοντελοποίηση και προσομείωση επιτρέπει τη βαθύτερη κατανόηση των μοριακών μηχανισμών ενός οργανισμού, συνεπώς ενός ζωντανού συστήματος. Ένα σύστημα αποτελεί ένα δίκτυο στοιχείων που αλληλεπιδρούν μεταξύ τους μέσω σχέσεων και ορίζουν ένα πολύπλοκα δομημένο σύνολο με συγκεκριμένα όρια. Μπορεί να είναι ανοιχτό ή κλειστό ανάλογα με το αν αλληλεπιδρά με το περιβάλλον του.

Στη μεταβολική μηχανική το υπό εξέταση σύστημα είναι το μεταβολικό δίκτυο βιοχημικών αντιδράσεων ενός οργανισμού και η μοντελοποίηση του δικτύου αυτού

οδηγεί στα Μοντέλα Γονιδιακής Κλίμακας (Genome-Scale Metabolic Reconstructions-GEMs).

Συγκεκριμένα αυτά τα μοντέλα συσχετίζουν το γονιδίωμα με τη μοριακή φυσιολογία. Κατακερματίζουν τα μεταβολικά μονοπάτια στις αντίστοιχες αντιδράσεις και ένζυμα και τα αναλύουν στο πλαίσιο του συνολικού δικτύου. Πιο απλοποιημένα, ένα μοντέλο συλλέγει όλες τις σχετικές μεταβολικές πληροφορίες ενός οργανισμού και τις συντάσσει σε ένα μαθηματικό μοντέλο. (Francke, et al., 2005)

Τα GEMs μπορούν να χωριστούν σε δύο κατηγορίες: τα Περιοριστικά Μοντέλα (ΠΜ) και τα Δυναμικά Μοντέλα (ΔΜ). Η πιο ευρέως χρησιμοποιούμενη κατηγορία είναι η πρώτη, λόγω απλότητας και των παραδοχών που μπορούν να γίνουν χωρίς να επηρεάζεται η λειτουργικότητα του μοντέλου. Τα δυναμικά μοντέλα περιέχουν εξισώσεις που περιγράφουν τους νόμους κινητικής, τις κινητικές παραμέτρους καθώς και τις αρχικές συγκεντρώσεις των μεταβολιτών.



Εικόνα 2.5 Περιοριστικά και Δυναμικά Μοντέλα. Η περιοχή λύσεων για το περιοριστικό μοντέλο περιγράφεται από το (a) Τυχαία δειγματοληψία μέσω αλγορίθμου hit-and-run και (b) γεωμετρική δειγματοληψία χρησιμοποιώντας τις γωνίες του κώνου ροών ως σημεία έναρξης. Ο χώρος λύσεων του δυναμικού μοντέλου περιγράφεται από το (c) μεταβολή τις αρχικής συγκέντρωσης μεταβολιτών και (d) τις κινητικές παραμέτρους. (e) Με περιορισμό των κινητικών παραμέτρων του δυναμικού μοντέλου μπορούμε να οριοθετήσουμε κινητικά δυνατές περιοχές ροών και να τις μεταφέρουμε στο περιοριστικό μοντέλο.

Αυτές οι δύο κατηγορίες μοντέλων αντιπροσωπεύουν τα δύο άκρα. Τα δυναμικά περιέχουν λεπτομερείς μηχανιστικές πληροφορίες που δίνουν λύσεις στιγμιαίας δυναμικής προσέγγισης στην ισοροπία σε οποιεσδήποτε αρχικές συνθήκες- γενικά συγκεντρώσεις ενζύμων και μεταβολιτών. Ακόμη δίνουν τη σταθερή κατάσταση που ορίζεται από τις συγκεντρώσεις μεταβολιτών. Οι ροές σταθερής κατάστασης υπολογίζονται από τις συγκεντρώσεις σταθερής κατάστασης και τους νόμους κινητικής. Τα περιοριστικά μοντέλα φαίνονται πιο μινιμαλιστικά. Δεν περιλαμβάνουν γνώση των χημικών αντιδράσεων πέραν της στοιχειομετρίας τους, οι λύσεις τους έχουν ροές σε σταθερή κατάσταση χωρίς πληροφορίες για συγκεντρώσεις. Αντί να δίνουν μία μοναδική λύση, παράγουν ένα υψηλής διάστασης συνεχές σταθερής κατάστασης λύσεων. (Machado, et al., 2012)

2.2 Διαδικασία μοντελοποίησης Γονιδιακής Κλίμακας-Γενικά

Υπάρχουν δύο τρόποι κατασκευής μεταβολικών μοντέλων:

- 1) Ακολουθώντας μία διαδικασία από κάτω προς τα πάνω (bottom up approach) με βάση γονιδιακές και βιβλιογραφικές πληροφορίες.
- 2) Ακολουθώντας την αντίστροφη διαδικασία από πάνω προς τα κάτω (top down approach)

Η πρώτη προσέγγιση οδηγεί στα μεταβολικά μοντέλα γονιδιακής κλίμακας, τα οποία είναι απολύτως λειτουργικά σε σύγκριση με τα μοντέλα που προκύπτουν από τη δεύτερη προσέγγιση και συνήθως υστερούν.

Συνολικά, η διαδικασία κατασκευής από κάτω προς τα πάνω αποτελείται από τέσσερα στάδια με επιμέρους βήματα το καθένα, τα οποία αναλύονται στο Κεφάλαιο 5. Είναι επαναληπτική και χρονοβόρα, διαρκεί από έξι μήνες έως και δύο χρόνια, απαιτώντας μια πληθώρα δεδομένων που πολλές φορές δεν είναι διαθέσιμα. Όσο περισσότερες πληροφορίες για τη φυσιολογία, τη βιοχημεία και τη γενετική του οργανισμού είναι διαθέσιμες τόσο πιο αξιόπιστο θα είναι το μοντέλο.

Οι βάσεις δεδομένων που αξιοποιούνται σε ολόκληρη τη διαδικασία μπορούν να χωριστούν στις εξής γενικές κατηγορίες:

1. Βιοχημείας. Πληροφορίες από εργαστηριακά πειράματα για την ύπαρξη και λειτουργία ενός ενζύμου σε έναν οργανισμό.
2. Γενομικής. Συσχετισμοί του γονιδιώματος ενός οργανισμού με τις λειτουργίες που επιτελεί.
3. Φυσιολογίας. Δομικές και λειτουργικές πληροφορίες.
4. In silico δεδομένα. Δεδομένα που έχουν προκύψει από τις *in silico* αναλύσεις και μαρτυρούν κενά του δικτύου που πρέπει να καλυφθούν.

Τα τέσσερα στάδια κατασκευής ενός μοντέλου γονιδιακής κλίμακας είναι τα εξής:

A. Δημιουργία προσχεδίου (draft reconstruction)

Αφορά την κατασκευή ενός προσχεδίου βασιζόμενο σε γονιδιακές και βιοχημικές βάσεις δεδομένων. Αποτελεί μια αρχική συλλογή υποψηφίων μεταβολικών λειτουργιών του οργανισμού που προκύπτουν άμεσα από το γονιδιώμα του. Μπορεί να αυτοματοποιηθεί μερικώς με κατάλληλα υπολογιστικά εργαλεία, αλλά δεν μπορεί να αντικατασταθεί εξ' ολοκλήρου από μία αυτοματοποιημένη διαδικασία καθώς απαιτείται η ανθρώπινη κρίση για την έγκυρη επιλογή και συλλογή των δεδομένων.

B. Μη αυτοματοποιημένη (χειροκίνητη) βελτίωση του προσχεδίου (manual refinement of reconstruction)

Το δεύτερο στάδιο είναι αυτό που απαιτεί τη μεγαλύτερη ακρίβεια καθώς το δίκτυο παίρνει την τελική του μορφή. Αν δεν είναι έγκυρη τότε ακυρώνεται η προβλεπτική ικανότητα του μοντέλου. Έτσι, επανεξετάζονται τα δεδομένα του πρώτου, προστίθενται καινούργια και συντίθενται ένα ολοκληρωμένο μεταβολικό δίκτυο. Οι πληροφορίες που αξιοποιούνται προέρχονται εκτός από τις βάσεις δεδομένων, από τη βιβλιογραφία και τις πειραματικές μετρήσεις. Αφορούν, πιο συγκεκριμένα δεδομένα, όπως τον τύπο των αντιδράσεων και το χώρο στον οποίο πραγματοποιούνται, τις συνθήκες κάτω από τις οποίες αυτές είναι εφικτές (στοιχειομετρία, pH, υποστρώματα, ενέργεια Gibbs κ.α.), τις σχέσεις γονιδίων-αντιδράσεων (GPR associations), κ.α. Δύο από τις σημαντικότερες αντιδράσεις που καθορίζονται εδώ είναι αυτή της βιομάζας (biomass reaction) και των ενεργειακών

απαιτήσεων ενός κυττάρου τόσο για την ανάπτυξη (Growth-associated ATP maintenance reaction-GAM) όσο και για τη συντήρηση του (Non-GAM reactions).

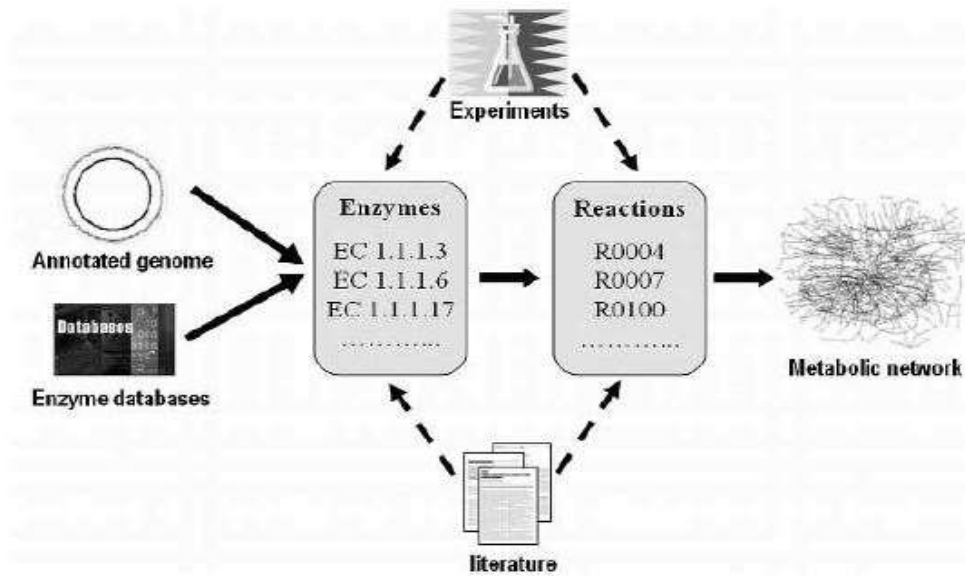
C. Μετατροπή του προσχεδίου σε μαθηματικό μοντέλο (conversion from reconstruction to mathematical model)

Αυτό το στάδιο αφορά τη μαθηματική αναπαράσταση του δικτύου και τον καθορισμό των ορίων του συστήματος που αυτό αντιπροσωπεύει. Αυτό το στάδιο είναι κυρίως αυτοματοποιημένο και καθορίζει το είδος της μοντελοποίησης που ακολουθείται καθώς η μαθηματική περιγραφή συνδέεται άμεσα με το σύστημα που εξετάζεται και τους σκοπούς της ανάλυσης. Συγκεκριμένα, αυτό το στάδιο μπορεί να οδηγήσει είτε σε περιοριστικά μοντέλα, είτε σε δυναμικά μοντέλα .

D. Αξιολόγηση του μοντέλου (network evaluation)

Στο τέταρτο και τελευταίο στάδιο γίνεται αξιολόγηση του μοντέλου *in silico* για τυχόν κενά που μπορεί να παρουσιάζει. Εάν υπάρχουν, καλύπτονται μέσω μιας επαναληπτικής διαδικασίας των τριών τελευταίων σταδίων. Τα πιο κρίσιμα ζητούμενα για να λήξει η επαναληπτική διαδικασία και να θεωρηθεί ολοκληρωμένο το μοντέλο, είναι να μπορεί να αναπαραχθεί και να επιβεβαιώνει τους φαινοτύπους που έχουν αποδειχθεί πειραματικά κάτω από διάφορες συνθήκες.

Σε ένα πέμπτο στάδιο, εφόσον το μοντέλο ανταποκρίνεται με επιτυχία σε όλα τα παραπάνω τότε μπορεί να λειτουργήσει και ως προβλεπτικό εργαλείο και διατίθεται στην επιστημονική κοινότητα.



Εικόνα 2.6 Κατασκευή μεταβολικού δικτύου από γονιδιακές πληροφορίες. (Ma & Zeng, 2006)

Όλα τα στάδια παρουσιάζουν τις δυσκολίες και τους περιορισμούς τους, γεγονός που επηρεάζει το τελικό αποτέλεσμα. Ιδιαίτερη σημασία πρέπει να δίνεται στη χρήση αξιόπιστων βάσεων δεδομένων και δη της γονιδιακής καθώς η κατασκευή ξεκινά από το γονιδίωμα κάθε οργανισμού. Οι σημαντικότερες αδυναμίες αφορούν την έλλειψη συγκεκριμένων βιολογικών δεδομένων των οργανισμών και μιας καθολικής, πρότυπης διαδικασίας κατασκευής που μπορεί να αυτοματοποιηθεί για να καλύψει τις ανάγκες μιας ραγδαίας αύξησης των γονιδιωμάτων που γίνονται γνωστά. (Thiele & Palsson, 2010)

2.3 Μαθηματική αναπαράσταση μοντέλων γονιδιακής κλίμακας

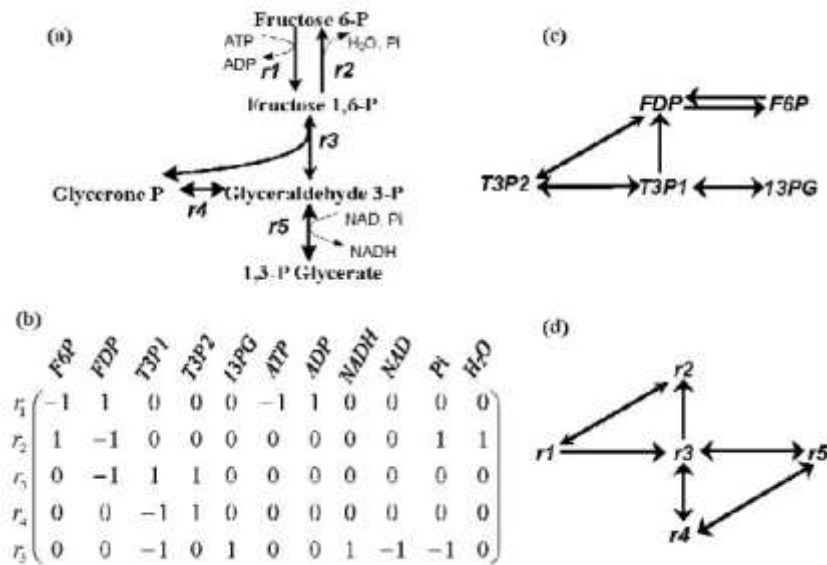
Η αναπαράσταση ενός συστήματος με μαθηματικούς φορμαλισμούς οδηγεί στα μαθηματικά/υπολογιστικά μοντέλα. Επιτρέπει την ακριβή και ξεκάθαρη ανάλυση τους μέσω της αναγωγής των μαθηματικών ιδιοτήτων σε συστημικές ιδιότητες και αντίστροφα.

Συγκεκριμένα, η μεταβολική μοντελοποίηση βασίζεται στην αναπαράσταση του δικτύου με τη μορφή ενός στοιχειομετρικού πίνακα S , την μετατροπή των βιολογικών περιορισμών, όπως η αρχή διατήρησης της μάζας, σε εξισώσεις και ανισότητες και τη λύση του συστήματος εξισώσεων που προκύπτει τελικά από τη σύνθεση των παραπάνω σε ένα μαθηματικό πρόβλημα. Η ερμηνεία των αποτελεσμάτων βασίζεται

στην αναγωγή των μαθηματικών ιδιοτήτων του πίνακα σε βιοχημικές/βιολογικές ιδιότητες του δικτύου. (Palsson, 2006)

2.3.1 Μέθοδοι μαθηματικής αναπαράστασης

Η σωστή μαθηματική αναπαράσταση ενός μεγάλου αριθμού αντιδράσεων οι οποίες έχουν συλλεχθεί για ένα συγκεκριμένο οργανισμό είναι απαραίτητη για την δομική ανάλυση των μεταβολικών δικτύων. Γενικά χρησιμοποιούνται δύο μέθοδοι: ο στοιχειομετρικός πίνακας και το γράφημα συνδεσιμότητας, όπως φαίνεται στην Εικ. 7. Στη στοιχειομετρική αναπαράσταση, οι σειρές και οι στήλες του στοιχειομετρικού πίνακα αναπαριστούν αντίστοιχα αντιδράσεις και μεταβολίτες. Μία θέση-κελί στον πίνακα με μη-μηδενική τιμή αντιπροσωπεύει το στοιχειομετρικό συντελεστή του αντίστοιχου μεταβολίτη στην αντίστοιχη αντίδραση. Μία θετική τιμή δηλώνει ότι είναι προϊόν, ενώ μια αρνητική τιμή ότι είναι υπόστρωμα (καταναλώνεται). Μηδενική τιμή δηλώνει ότι ο μεταβολίτης δε συμμετέχει στην αντίδραση.



Εικόνα 2.7 Μαθηματική αναπαράσταση των μεταβολικών δικτύων. (a) Το άνω κομμάτι του μονοπατιού της γλυκόλυσης, (b) ο στοιχειομετρικός πίνακας του μονοπατιού, (c) η αναπαράσταση με γράφημα των μεταβολιτών του μονοπατιού, (d) η αναπαράσταση με γράφημα των αντιδράσεων του μονοπατιού

Σε αντίθεση με τη στοιχειομετρική αναπαράσταση, η αναπαράσταση γραφήματος αποτελεί έναν απλοποιημένο τρόπο αναπαράστασης του μεταβολικού δικτύου. Όπως φαίνεται και στην Εικ. 7 δύο τύποι γραφήματος μπορούν να παραχθούν από ένα μεταβολικό δίκτυο: το γράφημα μεταβολιτών -όπου οι κόμβοι είναι μεταβολίτες και οι συνδέσεις αντιδράσεις- και με το γράφημα αντιδράσεων -όπου οι κόμβοι είναι αντιδράσεις και δύο αντιδράσεις συνδέονται αν ένας μεταβολίτης είναι το υπόστρωμα

της μίας αντίδρασης και το προϊόν της άλλης. Το γράφημα μεταβολιτών είναι σχετικά όμοιο με τον κλασικό τρόπο απεικόνισης των μεταβολικών δικτύων στη Βιοχημεία και γι' αυτό χρησιμοποιείται συχνά στη δομική ανάλυση των μεταβολικών δικτύων.

Λαμβάνοντας υπόψιν ότι πολλές αντιδράσεις είναι αναντίστροφες, πολλές συνδέσεις στο γράφημα έχουν κατεύθυνση. Σε σύγκριση με τη στοιχειομετρική αναπαράσταση, η αναπαράσταση γραφήματος είναι καταλληλότερη για απεικόνιση και ανάλυση δομής μεγάλης κλίμακας μεταβολικών δικτύων. Ωστόσο, αποτελεί απλοποιημένο τρόπο αναπαράστασης του δικτύου και «χάνει» πληροφορίες, όπως τους στοιχειομετρικούς συντελεστές των αντιδράσεων. (Ma & Zeng, 2006)

2.3.2 Βασικά στοιχεία του στοιχειομετρικού πίνακα S

Διαστάσεις S

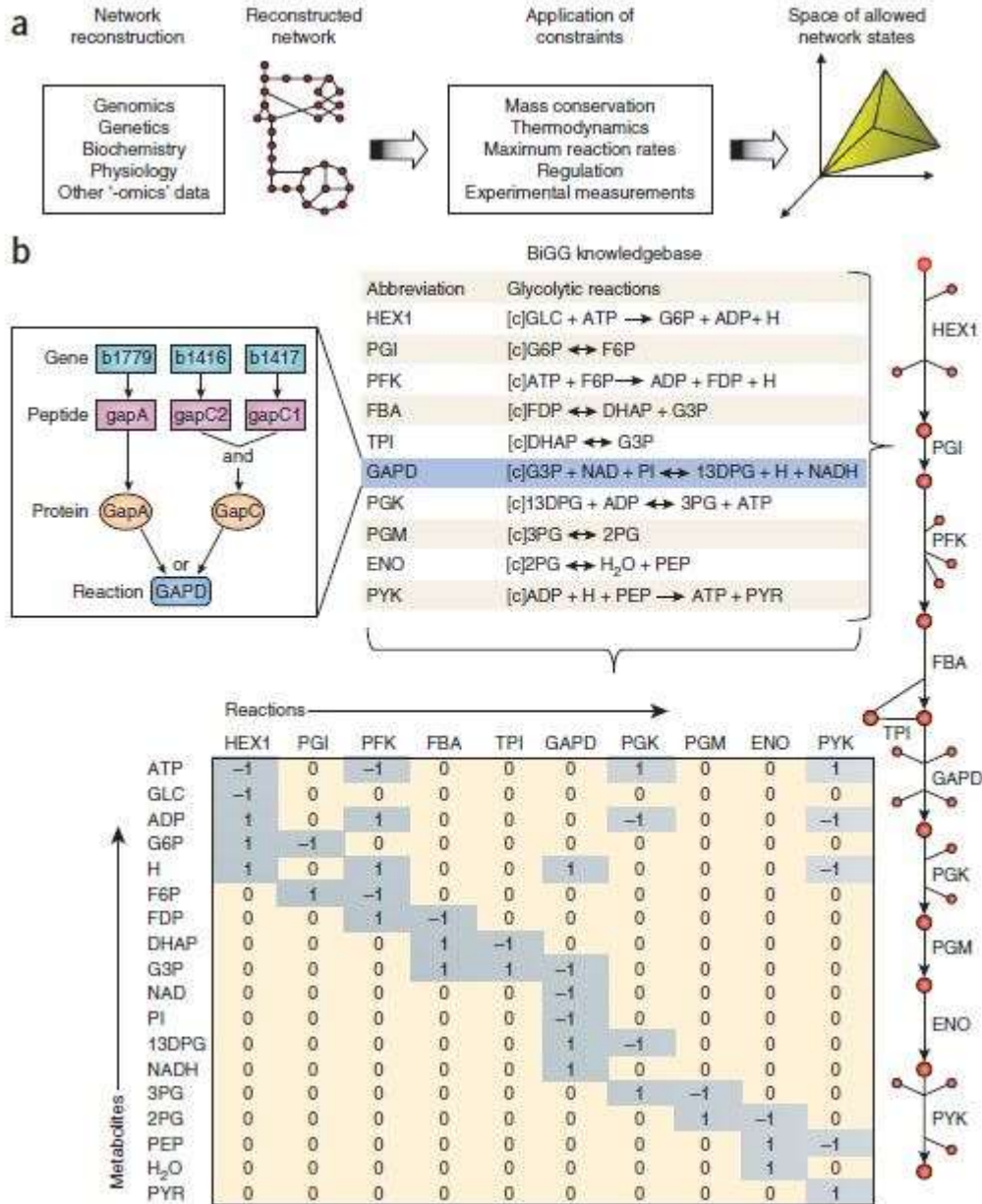
Ο S έχει διαστάσεις $m \times n$, όπου m οι μεταβολίτες και n οι αντιδράσεις του δικτύου. Κάθε γραμμή του πίνακα παριστάνει ένα μεταβολίτη και κάθε στήλη μία αντίδραση. Τα στοιχεία του πίνακα είναι οι στοιχειομετρικοί συντελεστές των μεταβολιτών της κάθε αντίδρασης που λαμβάνει μέρος στο δίκτυο. Είναι αρνητικοί για τα αντιδρώντα, θετικοί για τα προϊόντα και μηδενικοί για οποιονδήποτε μεταβολίτη δε συμμετέχει σε μια αντίδραση.

Οι ροές των αντιδράσεων του δικτύου παριστάνονται με ένα διάνυσμα v , μεγέθους n, ενώ οι συγκεντρώσεις των μεταβολιτών με ένα διάνυσμα x , μεγέθους m. Έτσι τα διανύσματα έχουν διαστάσεις:

$$\mathbf{dim}(x) = m \mathbf{dim}(v) = n \text{ και } \mathbf{dim}(S) = m \times n$$

Για ένα τυπικό μεταβολικό δίκτυο οι αντιδράσεις είναι περισσότερες από τους μεταβολίτες, δηλαδή $n > m$. Τότε ο πίνακας S είναι ανεπαρκούς τάξης και ισχύει $\text{Rank}(s) = r < m$.

Στην Εικ. 8 δίνεται η μετατροπή ενός απλού μεταβολικού μονοπατιού σε στοιχειομετρικό πίνακα.



Εικόνα 2.8 (α) Δημιουργία μεταβολικού δικτύου από ποικίλες πηγές βιολογικών δεδομένων. Οι δυνατότητες του μοντέλλου αξιολογούνται στο πλαίσιο φυσικών, χημικών, ρυθμιστικών περιορισμών. (β) Τα μοντέλα συνήθως κατασκευάζονται με τη βοήθεια της βάσης δεδομένων BiGG, που συνδέουν γονιδιακά δεδομένα με μεταβολική δραστηριότητα. Το μοντέλο μπορεί να μετατραπεί σε χάρτη (δεξιά) προς διευκόλυνση της απεικόνισης ή σε ένα φορμαλισμό μαθηματικής μοντελοποίησης, όπως είναι ο στοιχειομετρικός πίνακας (κάτω), ο οποίος μπορεί να χρησιμοποιηθεί για τη μελέτη ροής μάζας διαμέσου του δικτύου. (Schellenberger, et al., 2011)

Ο πίνακας S και τα ισοζύγια μάζας

Το μεγαλύτερο μέρος μιας *in silico* ανάλυσης βασίζεται στην εφαρμογή ισοζυγίων μάζας για κάθε μεταβολίτη. Τα ισοζύγια οδηγούν σε ένα σύστημα εξισώσεων η λύση του οποίου δίνει τις μεταβολικές ροές.

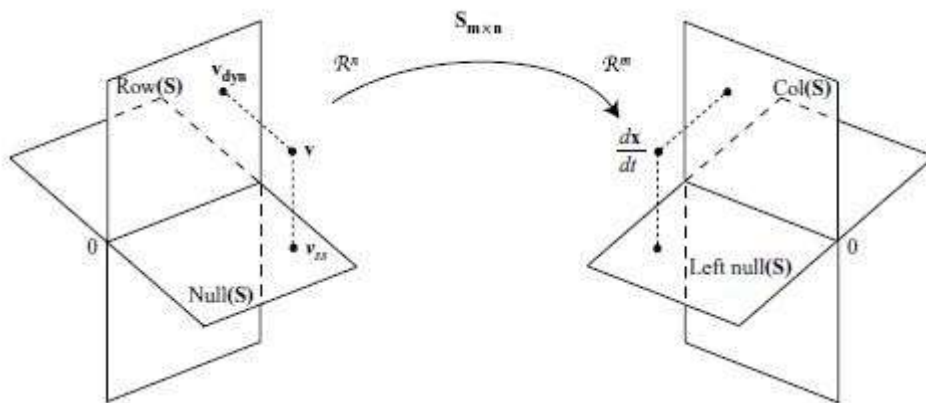
Κάθε σύστημα εξισώσεων μπορεί να αναπαρασταθεί με τη χρήση πινάκων και με τη σειρά τους οι πίνακες μπορούν να αποτελέσουν ένα γραμμικό μετασχηματισμό ανάμεσα σε δύο διανυσματικούς χώρους. Οι αναπαραστάσεις αυτές είναι σημαντικές διότι διευκολύνουν την επίλυση του συστήματος και συνεπώς την ανάλυση του μεταβολικού μοντέλου.

Παρακάτω παρουσιάζεται πως τα ισοζύγια μάζας συνδέονται με τον S ως γραμμικό μετασχηματισμό και τους τέσσερις γνωστούς υπόχωρους του.

Ο S ως γραμμικός μετασχηματισμός

Ο S μπορεί να θεωρηθεί ως ένας γραμμικός μετασχηματισμός του διανύσματος $v(x,k)=(v_1(x,k), v_2(x,k) \dots)$, που αντιπροσωπεύει τις ροές του δικτύου σε ένα διάνυσμα $x(t) =(x_1(t), x_2(t) \dots)$, που αντιπροσωπεύει τις συγκεντρώσεις των μεταβολιτών. Ο γραμμικός μετασχηματισμός αφορά τους υπόχωρους του πίνακα S. Ισχύει :

$$\frac{dx(t)}{dt} = S \cdot v(x, k) \quad [1]$$



Εικόνα 2.9 Ο πίνακας S ως γραμμικός μετασχηματισμός και οι τέσσερις θεμελιώδεις υποχώροι του (γραμμές, στήλες, μηδενοχώρος, αριστερός μηδενοχώρος)

Η εξίσωση [1] αναπαριστά το θεμελιώδες ισοζύγιο μάζας που χαρακτηρίζει όλες τις λειτουργικές καταστάσεις ενός μεταβολικού δικτύου. Αφορά τη δυναμική κατάσταση

του δικτύου, καθώς αφενός εμπεριέχει την μεταβολή των συγκεντρώσεων με το χρόνο (αριστερό μέλος) και αφετέρου την εξάρτηση των ροών από τις συγκεντρώσεις και τις κινητικές παραμέτρους (δεξί μέλος). Για κάθε μεταβολίτη ξεχωριστά γράφεται:

$$\frac{dx_i(t)}{dt} = \sum_k s_{ik} \cdot v_k(x_i, k_i) \quad [2]$$

όπου Σ είναι το άθροισμα όλων εκείνων των ροών που δημιουργούν και αποσυνθέτουν τον μεταβολίτη i στο χρονικό διάστημα dt . Η εφαρμογή της [2] για όλους τους μεταβολίτες ξεχωριστά οδηγεί στο μη γραμμικό σύστημα εξισώσεων που περιγράφει ένα δυναμικό μοντέλο.

Για να επιλυθεί το σύστημα [2] είναι απαραίτητη η γνώση κινητικών παραμέτρων οι οποίες πολλές φορές είναι άγνωστες και έτσι είναι αναγκαίες κάποιες ρεαλιστικές παραδοχές. Στα μεταβολικά δίκτυα, μία τέτοια παραδοχή είναι η υπόθεση της σταθερής κατάστασης η οποία τελικά μετατρέπει το δυναμικό σε περιοριστικό μοντέλο.

Η παραδοχή της σταθερής κατάστασης επιβάλλει στις συγκεντρώσεις των μεταβολιτών $x_i(t)$, να μην αλλάζουν με το χρόνο. Είναι λογική λόγω των πολύ μεγάλων διαφορών στους χρόνους αντίδρασης των ενζυμικά καταλυόμενων αντιδράσεων και των μεταβολών στους κυτταρικούς φαινοτύπους όπως είναι η κυτταρική ανάπτυξη. Έτσι, η σχέση [1] γίνεται:

$$\mathbf{S} \cdot \mathbf{v} = \mathbf{0} \quad [3]$$

Από την [3] προκύπτει ένα γραμμικό σύστημα εξισώσεων, η λύση του οποίου είναι πιο απλή καθώς δεν υπάρχει η παράμετρος του χρόνου άρα και η ανάγκη χρήσης των κινητικών παραμέτρων. Το σύστημα αυτό, που περιγράφει ένα περιοριστικό μοντέλο, μπορεί να είναι πλήρως ορισμένο (determined), υπερ-ορισμένο (overdetermined) ή υπό-ορισμένο (underdetermined).

Στα βιολογικά συστήματα είναι σχεδόν πάντα υπο-ορισμένο (underdetermined system) διότι δεν υπάρχουν αρκετές μετρημένες πειραματικά εξωτερικές ροές και έτσι οι εξισώσεις είναι λιγότερες απ' ότι οι άγνωστοι. Αυτό το σύστημα που έχει

πολλούς βαθμούς ελευθερίας για να λυθεί χρησιμοποιούνται υπολογιστικές μέθοδοι ανάλυσης υπό περιορισμούς όπως περιγράφεται στο Κεφάλαιο 3.

Υπόχωροι πίνακα

Από τον πίνακα S που ορίζει ένα σύνολο διανυσμάτων, δηλαδή ένα διανυσματικό χώρο, προκύπτουν τέσσερις θεμελιώδεις υπόχωροι. Για να οριστεί ένας υπόχωρος είτε πρέπει να είναι γνωστό ένα σύνολο διανυσμάτων που τον παράγουν, είτε κάποιες συνθήκες που πρέπει να ικανοποιούν τα διανύσματα που βρίσκονται μέσα σε αυτόν.

Το θεμελιώδες ισοζύγιο μάζας που παρουσιάστηκε παραπάνω αποτελεί μία συνθήκη που συνδέεται άρρηκτα με τους υπόχωρους ενός πίνακα. Μάλιστα, ο γραμμικός μετασχηματισμός παράγει τους υπόχωρους του πίνακα με αποτέλεσμα να συσχετίζει δυναμικές και σταθερές καταστάσεις όπως θα φανεί παρακάτω. Οι τέσσερις υπόχωροι είναι:

- *Χώρος Στηλών (ColumnSpace)*. Προκύπτει από το σύνολο των διανυσμάτων που ορίζουν οι στήλες του πίνακα S και έχει διαστάσεις $\dim(\text{Col}(S))=r$. Περιέχει όλες τις πιθανές χρονικές παραγώγους του διανύσματος των συγκεντρώσεων $x(t)$, (dx/dt) , και δείχνει τις θερμοδυναμικές επιρροές στις συγκεντρώσεις των μεταβολιτών του δικτύου (concentration states of the network).
- *Χώρος Γραμμών (RowSpace)*. Προκύπτει από το σύνολο των διανυσμάτων που ορίζουν οι γραμμές του S και έχει διαστάσεις $\dim(\text{Row}(S))=r$. Περιέχει το διανύσματα των ροών $v(x,k)$ δηλαδή όλες τις δυναμικές κατανομές ροών των αντιδράσεων και δείχνει τις θερμοδυναμικές επιρροές στο ρυθμό των αντιδράσεων.
- *Μηδενόχωρος (NullSpace)*. Προκύπτει από τη συνθήκη $S \cdot v=0$ που ορίζει ένα ομογενές σύστημα εξισώσεων, οι λύσεις του οποίου αποτελούν τον υπόχωρο $\dim(\text{Null}(S))=n-r$. Στα μεταβολικά συστήματα η συνθήκη αυτή αντιπροσωπεύει τη σταθερή κατάσταση και έτσι δίνει το σύνολο των κατανομών των ροών υπό σταθερή κατάσταση.
- *Αριστερός μηδενόχωρος (LeftNull)*. Προκύπτει από τη συνθήκη $L \cdot S=0$ και έχει διαστάσεις $\dim(\text{LeftNull}(S))=m-r$. Περιγράφει όλες τις σχέσεις συντήρησης (μεταβλητές που δεν αλλάζουν με το χρόνο) σε ένα μεταβολικό

δίκτυο, για παράδειγμα το άθροισμα των μεταβολιτών που δεν αλλάζουν με το χρόνο.

Από τα παραπάνω συνάγεται ότι οι δύο πρώτοι υπόχωροι αφορούν συστήματα αφενός σε μη σταθερή κατάσταση και αφετέρου σε μη θερμοδυναμική ισορροπία (δυναμικά συστήματα). Η ανάλυση τους απαιτεί τη γνώση κινητικών παραμέτρων και είναι πολύπλοκη σε μεγάλης κλίμακας μεταβολικά δίκτυα. Αντίθετα, οι δύο τελευταίοι αναφέρονται στην σταθερή κατάσταση και η ανάλυση τους είναι πιο απλή. Μέχρι στιγμής, λόγω της απλότητας, η ανάλυση των υπόχωρων ενός πίνακα έχει περιοριστεί σε αυτούς που αναφέρονται σε σταθερές καταστάσεις. Αν και δυναμικές προσομοιώσεις έχουν υπάρξει αρκετές, ακόμα δεν έχει εδραιωθεί η σχέση τους και η ανάλυση τους μέσω των υπόχωρων.

Συμπερασματικά, ο κάθε υπόχωρος παρέχει διαφορετικές πληροφορίες για ένα μεταβολικό δίκτυο και μόνο η ανάλυση όλων των υποχώρων συνδυαστικά μπορεί να οδηγήσει σε πιο ρεαλιστικές προσομοιώσεις. Τα φυσικά συστήματα μπορούν να απλοποιηθούν μέσα από μία σειρά παραδοχών αλλά δεν μπορεί να αγνοηθεί η περιπλοκότητα της φύσης και ο δυναμικός χαρακτήρας τους. Έτσι, η ανάλυση μέσω του πίνακα S και των ιδιοτήτων των υπόχωρων του, προσφέρουν μία σημαντική βάση διερεύνησης της συμπεριφοράς του συστήματος.

Γραμμική εξάρτηση και ερμηνεία των αποτελεσμάτων

Μία σημαντική έννοια που σχετίζεται με τον πίνακα S και συνεπώς με τους διανυσματικούς υπόχωρους είναι η γραμμική εξάρτηση. Γραμμικά εξαρτημένα διανύσματα ονομάζονται αυτά που προκύπτουν ως γραμμικός συνδυασμός άλλων διανυσμάτων. Δεν μεγαλώνουν τον διανυσματικό χώρο με αποτέλεσμα να είναι περιττά. Αντίστοιχα, γραμμικά εξαρτημένο είναι ένα σύστημα οι εξισώσεις του οποίου προκύπτουν από άλλες και έτσι δεν προσφέρουν κάποιον επιπλέον περιορισμό.

Ο εντοπισμός των γραμμικά εξαρτημένων διανυσμάτων ή εξισώσεων και η αφαίρεση τους από το σύστημα έχει τεράστια σημασία, τόσο μαθηματική όσο και βιολογική. Αφενός, από τη μαθηματική σκοπιά το πρόβλημα απλουστεύεται γιατί μειώνονται οι

εξισώσεις και ο χώρος των λύσεων διευκολύνοντας τους υπολογισμούς και την ανάλυση. Αφετέρου, από βιολογικής σκοπιάς η εξάλειψη της γραμμικής εξάρτησης οδηγεί σε ένα απλουστευμένο μεταβολικό δίκτυο του οποίου οι σχέσεις μπορούν να αναδειχθούν και να ερμηνευτούν πιο εύκολα.

Ένα ακόμα σημαντικό στοιχείο για την απλούστευση της ερμηνείας των αποτελεσμάτων, αλλά και για την κάλυψη όλου του χώρου λύσεων και όχι μόνο μίας βέλτιστης ή άλλης λύσης, είναι η επιλογή της βάσης του διανυσματικού χώρου και των υποχώρων του. Μια κατάλληλη βάση αποτελείται από ένα σύνολο γραμμικά ανεξάρτητων διανυσμάτων και περιγράφει όλο το χώρο αποκαλύπτοντας σημαντικά βιολογικά στοιχεία όπως τις τοπολογικές ιδιότητες ενός δικτύου. Για γραμμικούς χώρους συνηθέστερες βάσεις είναι οι γραμμικές/ορθοκανονικές και μία χαρακτηριστική μέθοδος εύρεσης είναι η ανάλυση πίνακα σε ιδιάζουσες τιμές (SingularValueDecomposition).

Μορφές του S

Ο πίνακας μπορεί να πάρει διάφορες μορφές εξυπηρετώντας κάθε φορά το ζητούμενο του προβλήματος και τα όρια που επιβάλλει αυτό στο σύστημα.

Τα όρια αυτά δημιουργούν νοητά ένα εσωτερικό και ένα εξωτερικό χώρο στο κύτταρο και τα μέρη του, χωρίζοντας αντίστοιχα και τις αντιδράσεις σε εσωτερικές που συμβολίζονται με το διάνυσμα v_i και ανταλλαγής, b_i . Όμοια και για τους μεταβολίτες όπου x_i , το διάνυσμα των εσωτερικών και c_i , των εξωτερικών. Έτσι, προκύπτει ο συνολικός στοιχειομετρικός πίνακας S.

$$S_{tot} : \begin{matrix} & & v_i & & b_i \\ & x_i & & & \\ & c_i & & & \\ & & & 0 & \end{matrix} \left(\begin{array}{c|c} & \\ \hline & \end{array} \right)$$

Εικόνα 2.10 Ο συνολικός στοιχειομετρικός πίνακας.

Από το συνολικό πίνακα, ανάλογα με το ζητούμενο, μπορούμε να πάρουμε κάποιους υποπίνακες. Αν δεν λάβουμε υπόψιν τις εξωτερικές ροές, c_i προκύπτει ο υποπίνακας

ανταλλαγής, S_{exch} , ο οποίος περιέχει μόνο τις εσωτερικές ροές και τις ροές ανταλλαγής με το περιβάλλον. Ο εσωτερικός υποπίνακας προκύπτει αν θεωρήσουμε το κύτταρο ως κλειστό σύστημα, οπότε μπορούμε να εστιάσουμε στις εσωτερικές ροές και να μελετήσουμε τις ιδιότητες του. (Palsson, 2006)

$$S_{\text{exch}} : x_j \begin{pmatrix} v_i & b_j \\ \vdots & \vdots \\ \vdots & \vdots \end{pmatrix} \quad S_{\text{int}} : x_j \begin{pmatrix} v_i \\ \vdots \\ \vdots \end{pmatrix}$$

Εικόνα 2.11 Ο στοιχειομετρικός πίνακας ανταλλαγής S_{exch} και ο εσωτερικός πίνακας S_{int} . (Palsson, 2006)

Ακόμα μια χρήσιμη μορφή του πίνακα, που διευκολύνει την εύρεση των τοπολογικών ιδιοτήτων του, είναι η δυαδική μορφή (binary form). Ο S παίρνει τη καινούργια μορφή S' , για τα στοιχεία του οποίου ισχύει:

$$s'_{ij}=0 \text{ if } s_{ij}=0$$

$$s'_{ij}=1 \text{ if } s_{ij} \neq 0$$

Έτσι ο πίνακας S' , αποτελείται μόνο από μηδέν και ένα. Εάν το στοιχείο s'_{ij} είναι μονάδα τότε σημαίνει ότι ο μεταβολίτης i λαμβάνει μέρος στην αντίδραση j .

Ο S ως πίνακας σύνδεσης ενός γραφήματος

Κάθε δίκτυο αναπαριστάται με ένα γράφημα που αποτελείται από κόμβους και συνδέσεις μεταξύ των κόμβων, όπως αναφέρθηκε και πιο πάνω. Στο μεταβολικό δίκτυο οι κόμβοι είναι οι μεταβολίτες και οι σχέσεις οι αντιδράσεις. Αντίστοιχα, στον πίνακα οι αντιδράσεις αποτελούν τις στήλες και οι μεταβολίτες τις γραμμές. Έτσι, ο πίνακας S αναπαριστά και αυτός το ίδιο γράφημα, που ονομάζεται χάρτης αντιδράσεων και λειτουργεί ως πίνακας σύνδεσης του χάρτη αυτού.

Δύο αριθμοί που αποτελούν βασικά τοπολογικά χαρακτηριστικά του πίνακα S , είναι ο αριθμός συμμετοχής (participatio nnumber) μιας αντίδρασης και ο αριθμός

συνδεσιμότητας (connectivity number) ενός μεταβολίτη. Για να υπολογιστούν πρέπει ο πίνακας S να βρίσκεται σε δυαδική μορφή.

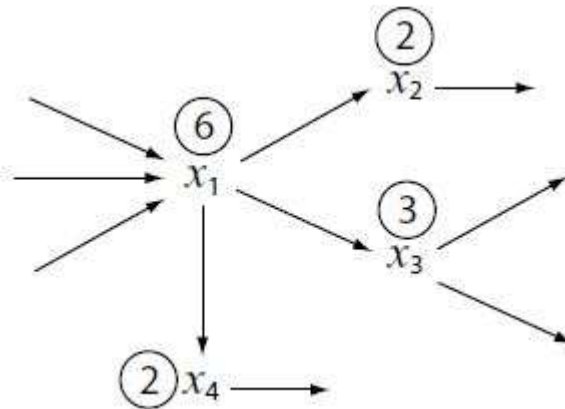
Ο αριθμός συμμετοχής, ορίζει σχέσεις μεταβολιτών σε μία αντίδραση και αποτελεί το άθροισμα όλων των μη μηδενικών στοιχείων μιας στήλης του πίνακα. Ορίζεται ως

$$\pi_j = \sum s_{ij} \quad [4]$$

ενώ ο αριθμός συνδεσιμότητας, ορίζει σχέσεις αντιδράσεων στις οποίες συμμετέχει ένας μεταβολίτης και αποτελεί το άθροισμα των μη μηδενικών στοιχείων μιας γραμμής. Ορίζεται ως

$$\rho_i = \sum s_{ij} \quad [5]$$

Στην Εικ.12 φαίνεται ένα απλό παράδειγμα συνδεσιμότητας των κόμβων σε ένα δίκτυο.



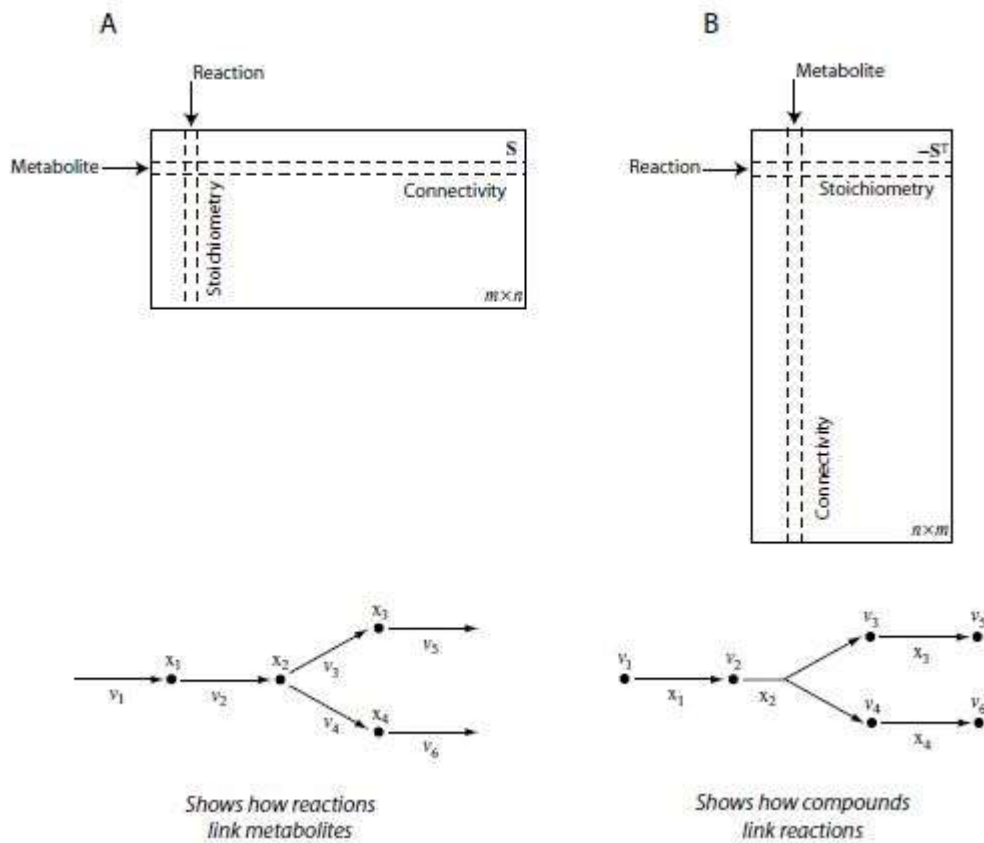
Εικόνα 2.12 Ένα απλό δίκτυο όπου παρουσιάζονται οι σχέσεις των κόμβων του δικτύου. Πάνω από κάθε κόμβο απεικονίζεται ο αριθμός συνδεσιμότητας.

Η φύση των συνδέσεων που ορίζουν οι δύο παραπάνω σχέσεις είναι διαφορετική. Ο αριθμός συμμετοχής ορίζει δεσμευτικές συνδέσεις (hard connections), ενώ ο αριθμός συνδεσιμότητας χαλαρές συνδέσεις (soft connections). Οι συνδέσεις του πρώτου τύπου είναι αναγκαίες και μη διαπραγματεύσιμες ενώ του δεύτερου το αντίθετο. Αυτό εξηγείται από το γεγονός ότι δεν μπορεί να συμβεί μια αντίδραση χωρίς να συμμετέχουν όλοι οι μεταβολίτες της, ενώ μπορεί ένας μεταβολίτης να μη συμμετέχει ταυτόχρονα σε όλες τις αντιδράσεις με τις οποίες συνδέεται.

Οι παραπάνω σχέσεις είναι σημαντικές γιατί αποκαλύπτουν στοιχεία για τη φύση των μεταβολιτών ενός δικτύου. Για παράδειγμα ένας μεταβολίτης με υψηλή

συνδεσιμότητα ίσως επηρεάζει περισσότερο τους φαινοτύπους ενός οργανισμού, καθώς συμμετέχει σε μία πληθώρα αντιδράσεων και έτσι μπορεί να αποτελέσει στόχο για την έρευνα φαρμάκων σε παθολογικές καταστάσεις. Ωστόσο, οι σχέσεις αυτές θα πρέπει πάντα να εξετάζονται σε συνάρτηση με τους επιτρεπτούς φαινοτύπους και να υπακούουν στη βιολογία του οργανισμού.

Στην Εικ.13 φαίνεται ο πίνακας S και ο ανάστροφός του καθώς και πως συνδέονται μεταβολίτες και αντιδράσεις. Ο ανάστροφος του πίνακα S παρόλο που χρησιμοποιείται σπανιότερα αναδεικνύει ανάλογες σχέσεις με αυτές του S . (Palsson, 2006)



Εικόνα 2.13 Ο πίνακας S και ο αντίστροφος του, ως πίνακες συνδεσιμότητας (α) στον πίνακα S φαίνεται πως οι αντιδράσεις συνδέουν τους μεταβολίτες και (β) στον αντίστροφο $-S^T$ φαίνεται πως οι μεταβολίτες συνδέουν τις αντιδράσεις.

2.4 Βιβλιογραφία

Francke, C., Siezen, R. J. & Teusink, B., 2005. Reconstructing the metabolic network of a bacterium from its genome. *Trends in Microbiology*, 13(11), pp. 550-558.

Lee, S. Y. et al., 2010. Systems Biology, Genome-Scale Models, and Metabolic Engineering. In: C. N. Smolke, ed. *The Metabolic Pathway Engineering Handbook: Tools and Applications*. Boca Raton: CRC Press, pp. 15-5-15-6.

Machado, D. et al., 2012. Exploring the gap between dynamic and constraint-based models of metabolism. *Metabolic Engineering*, 14(1), pp. 112-119.

Ma, H.-W. & Zeng, A.-P., 2006. Reconstruction of Metabolic Networks from Genome Information and Its Structural and Functional Analysis. In: A. Kriete & R. Eils, eds. *Computational Systems Biology*. 1st ed. Burlington: Elsevier Academic Press, pp. 169-175.

Palsson, B. . Ø., 2006. Mathematical Representation of Reconstructed Networks. In: *Systems Biology (Properties of Reconstructed Networks)*. New York: Cambridge University Press, pp. 87-178.

Schellenberger, J. et al., 2011. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0.. *Nature Protocols*, 6(9), pp. 1290-1307.

Thiele, I. & Palsson, B. . Ø., 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, 5(1), pp. 93-121.

ΚΕΦΑΛΑΙΟ 3

ΑΝΑΛΥΣΗ ΜΟΝΤΕΛΩΝ ΓΟΝΙΔΙΑΚΗΣ ΚΛΙΜΑΚΑΣ ΥΠΟ ΠΕΡΙΟΡΙΣΜΟΥΣ

Η θεωρία της εξέλιξης είναι θεμελιώδης για τις επιστήμες της βιολογίας. Επί της ουσίας, για να επιβιώσει ένας οργανισμός σε ένα περιβάλλον με περιορισμένους πόρους, θα πρέπει να ικανοποιούνται εκατοντάδες περιορισμοί, οι οποίοι ελαχιστοποιούν τον αριθμό πιθανών φαινότυπων. Όλοι οι φαινότυποι που εκφράζονται θα πρέπει να ικανοποιούν τους περιορισμούς που επιβάλλονται από τις μοριακές λειτουργίες του κυττάρου. Όλες οι διεργασίες του ζωντανού οργανισμού θα πρέπει να συμμορφώνονται με τους φυσικούς νόμους, όπως τη διατήρηση μάζας και ενέργειας. Επομένως η αναγνώριση και ο καθορισμός των περιορισμών για τον ορισμό του εύρους των επιτρεπόμενων φαινότυπων αποτελεί μια ουσιαστική προσέγγιση στην κατανόηση των βιολογικών συστημάτων, της λειτουργίας και εξέλιξης των οργανισμών.

Έχοντας κατασκευάσει επιτυχώς ένα μοντέλο γονιδιακής κλίμακας μπορούμε να προσδιορίσουμε εκείνους τους περιορισμούς κάτω από τους οποίους λειτουργούν αυτά τα δίκτυα. Ένα πλαίσιο για την κατασκευή και ανάλυση υπό περιορισμούς (Constraint-based reconstruction and analysis-COBRA) εφαρμόζεται για τη μελέτη των πιθανών φαινότυπων που προκύπτουν από ένα γονιδίωμα. Διάφορες *in silico* μέθοδοι στοχεύουν στον προσδιορισμό των ικανοτήτων και χαρακτηριστικών των μικροοργανισμών. Αυτές οι μέθοδοι διαμορφώνουν τη βάση της *in silico* ανάλυσης των οργανισμών. Σε αυτό το κεφάλαιο παρουσιάζονται οι μέθοδοι που έχουν αναπτυχθεί στο πλαίσιο COBRA.

3.1 Μέθοδοι αναλύσης υπό περιορισμούς

3.1.1 Περιορισμοί στις κυτταρικές λειτουργίες

Παρακάτω περιγράφονται σε τέσσερις κατηγορίες:

- A. Φυσικοχημικοί περιορισμοί.** Η μάζα, η ενέργεια και η ορμή πρέπει να διατηρούνται. Τα περιεχόμενα ενός κυττάρου είναι πυκνά τοποθετημένα και σχηματίζουν ένα περιβάλλον όπου το ιξώδες μπορεί να είναι 100-1000 φορές μεγαλύτερο του νερού. Οι ρυθμοί διάχυσης των μακρομορίων μέσα σε ένα κύτταρο είναι γενικά χαμηλοί και περιοριστικοί, ανάλογα το μέγεθος του μορίου. Ο εγκλωβισμός πολλών μορίων σε ημιπερατές μεμβράνες προκαλεί υψηλή ωσμωτικότητα γι'αυτό και τα κύτταρα απαιτούν μηχανισμούς για να αντιμετωπίσουν την ωσμωτική πίεση (όπως νάτριο-κάλιο ή κυτταρικό τοίχωμα). Οι ρυθμοί αντιδράσεων προσδιορίζονται από τοπικές συγκεντρώσεις μέσα στα κύτταρα και μπορούν να περιοριστούν από τη μεταφορά μάζας. Οι αριθμοί μετάβασης ενζύμων είναι γενικά μικρότεροι από 10^4 sec^{-1} και το ανώτατο όριο των ρυθμών αντιδράσεων είναι ίσι με τον αριθμό μετάβασης επί τη συγκέντρωση του ενζύμου. Οι βιοχημικές αντιδράσεις θα πρέπει να έχουν αρνητική μεταβολή ενέργειας Gibbs για να ακολουθούν κατεύθυνση προς τα εμπρός.
- B. Τοποβιολογικοί περιορισμοί.** Ο συνωστισμός των μορίων μέσα στο κύτταρο οδηγεί σε τοποβιολογικούς, ή αλλιώς τριών διαστάσεων περιορισμούς που επηρεάζουν τόσο τη μορφή όσο και τη λειτουργία των βιολογικών συστημάτων. Για παράδειγμα, σε pH 7.6, ο *E.coli* τυπικά θα περιέχει περίπου 16 ιόντα υδρογόνου. Αυτός ο εξαιρετικά χαμηλός αριθμός υποδηλώνει ότι μπορεί να μην έχει νόημα να μελετάμε το εσωκυτταρικό pH ενός κυττάρου από την άποψη «χονδρικών» μέσων όρων και ότι η παρακολούθηση των ιόντων υδρογόνου είναι κρίσιμη στο πλαίσιο των μοντέλων γονιδιακής κλίμακας.

C. Περιβαλλοντικοί περιορισμοί. Οι περιβαλλοντικοί περιορισμοί σε ένα κύτταρο είναι εξαρτώμενοι χρόνου και συνθηκών. Παραδείγματα: pH, θερμοκρασία, ωσμωτικότητα και διαθεσιμότητα δεκτών ηλεκτρονίων. Αυτοί οι περιορισμοί είναι θεμελιώδους σημασίας για την ποσοτική ανάλυση των μικροοργανισμών. Ο καθορισμός του μέσου ανάπτυξης και ορθώς καταγεγραμμένων περιβαλλοντικών συνθηκών απαιτούνται για την ενσωμάτωση δεδομένων από εργαστήρια στα ποσοτικά μοντέλα τα οποία είναι εξίσου περιγραφικά και προβλεπτικά.

D. Ρυθμιστικοί περιορισμοί. Οι ρυθμιστικοί περιορισμοί διαφέρουν από τις τρεις παραπάνω κατηγορίες καθώς είναι αυτοεπιβαλλόμενοι και υπόκεινται σε εξελικτική αλλαγή. Επιτρέπουν στο κύτταρο να εξαλείφει φαινότυπους οι οποίοι είναι μερικώς βέλτιστοι και να περιορίζεται σε συμπεριφορές αυξημένης ανάπτυξης.

3.1.2 Μαθηματική αναπαράσταση των περιορισμών

Μετά την αναγνώριση και τον καθορισμό τους, οι περιορισμοί πρέπει να περιγραφούν μαθηματικά. Με τη μαθηματική τους μορφή μπορούν να χρησιμοποιηθούν στην ανάλυση *in silico*.

Οι περιορισμοί μπορούν να κατηγοριοποιηθούν γενικά είτε ως ισοζύγια είτε ως όρια. Τα ισοζύγια είναι περιορισμοί που σχετίζονται με συντηρούμενα ποσά, όπως ενέργεια, μάζα, οξειδοαναγωγικό δυναμικό και ορμή κ.α.. Τα όρια είναι περιορισμοί που μειώνουν το αριθμητικό εύρος μεμονωμένων μεταβλητών και παραμέτρων όπως συγκεντρώσεις, ροές ή κινητικές σταθερές.

Η διατήρηση της μάζας είναι ένα παράδειγμα ισοζυγίου. Σε σταθερή κατάσταση δεν υπάρχει συσσώρευση ή εξάντληση μεταβολιτών στο δίκτυο, οπότε ο ρυθμός παραγωγής κάθε μεταβολίτη θα πρέπει να ισούται με το ρυθμό κατανάλωσής του. Αυτό το ισοζύγιο ροών μπορεί να παρασταθεί μαθηματικά ως:

$$S \cdot v = 0$$

Όπου v είναι ένα διάνυσμα ροών μέσα στο μεταβολικό δίκτυο και S είναι ο στοιχειομετρικός πίνακας που περιέχει τη στοιχειομετρία των αντιδράσεων του

μοντέλου. Αντίστοιχα ισοζύγια μπορούν να γραφούν για την ωσμωτική πίεση, την ελεύθερη ενέργεια γύρω από βιοχημικούς βρόγχους.

Μπορούν ακόμα να προσδιοριστούν τα όρια που περιορίζουν περαιτέρω τις τιμές μεμονωμένων μεταβλητών, δηλαδή συγκεντρώσεων, ροών και κινητικών σταθερών. Εφαρμόζονται άνω και κάτω όρια στις μεμονωμένες ροές:

$$v_{\min} \leq v \leq v_{\max}$$

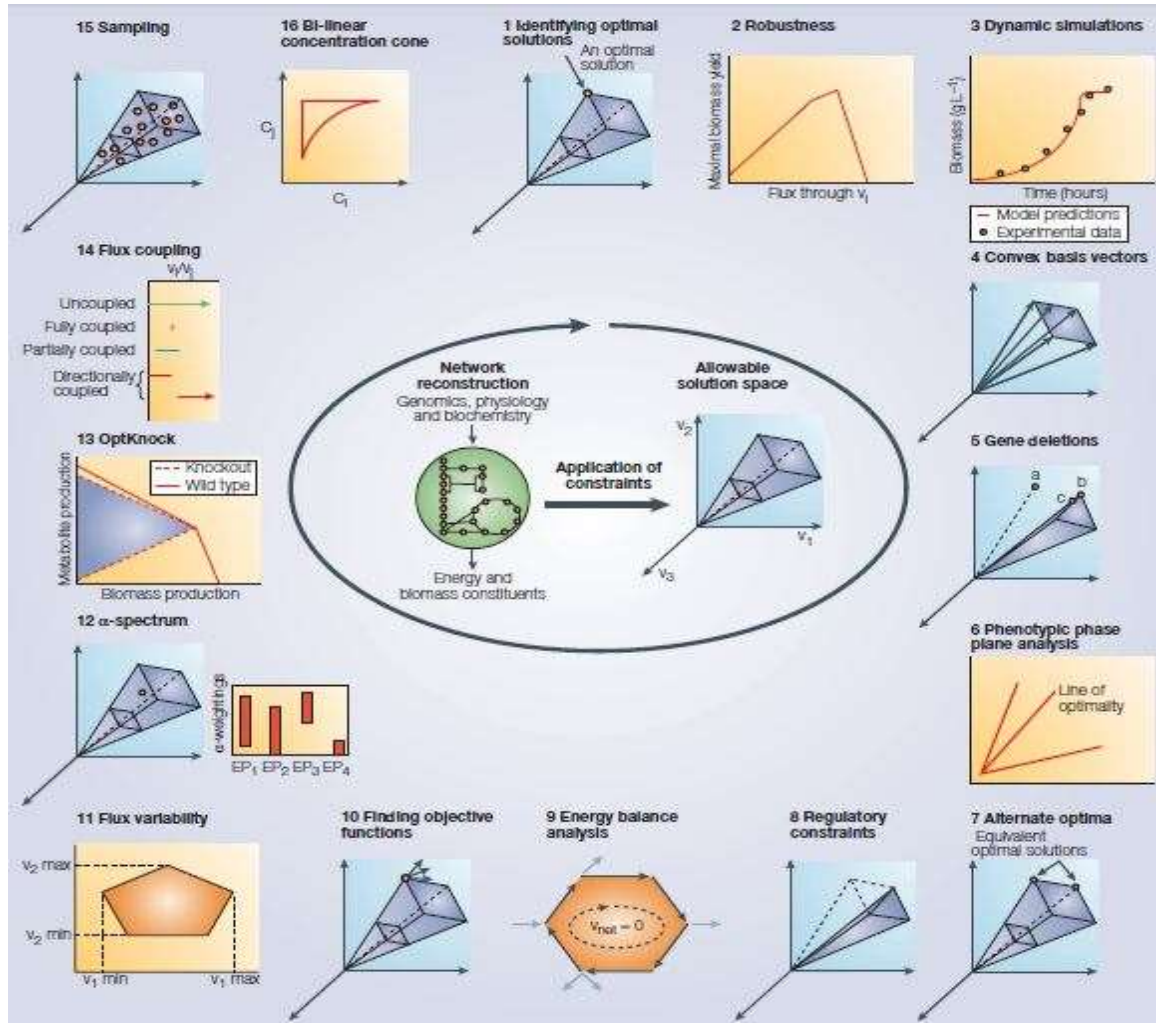
Για στοιχειακές (και αναντίστρεπτες) αντιδράσεις ισχύει:

$$v_{\min} = 0$$

Επιβάλλονται συγκεκριμένα άνω όρια (v_{\max}) στις αντιδράσεις, τα οποία βασίζονται σε μετρήσεις χωρητικότητας των ενζύμων. Οι συγκεντρώσεις πρέπει πάντα να έχουν θετική τιμή, αυτός ο περιορισμός θέτει κάτω όριο στις τιμές συγκέντρωσης. Οι κινητικές σταθερές περιορίζονται να έχουν θετική τιμή και άνω όριο το οποίο βασίζεται στη συχνότητα σύγκρουσης ($0 \leq k \leq k_{\max}$).

3.1.3 Εργαλεία ανάλυσης δικτύων

Με το συνδυασμό περιορισμών μέσω ισοζυγίων και ορίων, μειώνονται οι επιτρεπόμενες λειτουργικές καταστάσεις των κατασκευασμένων μοντέλων. Σε μαθηματικούς όρους, το εύρος των επιτρεπτών καταστάσεων του δικτύου περιγράφεται από ένα χώρο λύσεων που αναπαριστά τον δυνητικό φαινότυπο ενός οργανισμού. Εάν τα ισοζύγια και τα όρια περιγράφονται από γραμμικές εξισώσεις, τότε ο χώρος λύσεων είναι ένα πολύτοπο σε μεγάλων διαστάσεων χώρο. Αν οι περιορισμοί είναι διγραμμικοί, όπως αυτοί που προκύπτουν από κινητικές μάζας στοιχειακών αντιδράσεων, ο χώρος λύσεων είναι κοίλος.



Εικόνα 3.14 Ένα αναπτυσσόμενο εργαλείο για την ανάλυση υπό περιορισμούς. Τα δύο βήματα που απαιτούνται για τη διαμόρφωση του χώρου λύσεων είναι η κατασκευή του μοντέλου και η επιβολή των περιορισμών-φαινόνται στο κέντρο της εικόνας. Παρουσιάζονται οι μέθοδοι ανάλυσης υπό περιορισμούς.

Η επιβολή περιορισμών σε ένα κατασκευασμένο δίκτυο γονιδιακής κλίμακας (GENRE) οδηγεί σε ένα μοντέλο γονιδιακής κλίμακας (GEM) ενός οργανισμού, το οποίο μπορεί να μελετηθεί ως προς τις ικανότητες και τα χαρακτηριστικά του.

Οι μέθοδοι COBRA αποτελούνται από δύο βασικά βήματα:

1^ο : Κατασκευή του Μεταβολικού Μοντέλου Γονιδιωματικής Κλίμακας (Genome Scale Network Reconstruction-GENRE)

2^ο : Εφαρμογή κατάλληλων περιορισμών in silico στο ΜΜΓΚ ώστε να προκύψει ένα περιοριστικό μοντέλο (Genome Scale Model-GEM)

Η πληθώρα των μεθόδων COBRA χωρίζεται στις εξής κατηγορίες:

- Εύρεση καλύτερης/ βέλτιστης κατάστασης στο επιτρεπτό εύρος.
- Διευρεύνηση εξάρτησης ροών.
- Μελέτη όλων των επιτρεπτών καταστάσεων.
- Αλλαγή πιθανών φαινότυπων ως αποτέλεσμα γενετικών παρεκκλίσεων.
- Καθορισμός και επιβολή περαιτέρω περιορισμών.

Βέλτιστες καταστάσεις

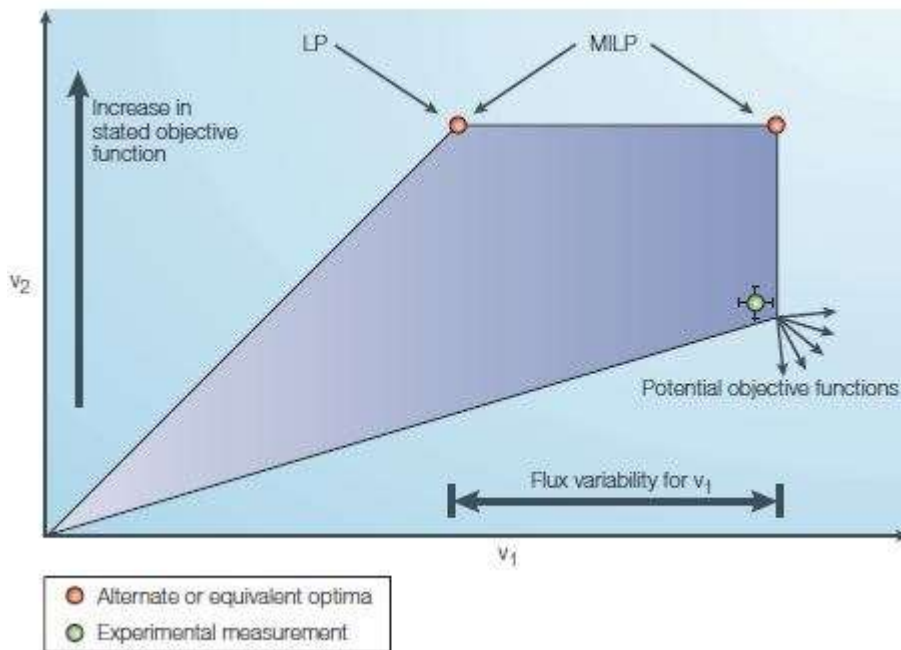
Ο μαθηματικός προγραμματισμός μπορεί να χρησιμοποιηθεί στην αναγνώριση καταστάσεων δικτύων βιοχημικών αντιδράσεων, οι οποίες μεγιστοποιούν μία συγκεκριμένη λειτουργία του δικτύου. Αυτή η επιθυμητή λειτουργία ορίζεται και περιγράφεται μαθηματικά, παίρνοντας τη μορφή αντικειμενικής συνάρτησης (Z). Οι αντικειμενικές συναρτήσεις δημιουργούνται για τρεις κυρίως λόγους: πρώτον, την εξερεύνηση φαινοτυπικών δυνατοτήτων ενός GENRE, δεύτερον, τον καθορισμό πιθανών φυσιολογικών καταστάσεων με επιλογή αντικειμενικής συνάρτησης που αντιπροσωπεύει πιθανές φυσιολογικές λειτουργίες, όπως η βιομάζα ή η παραγωγή ATP και τρίτον, ο σχεδιασμός στελεχών στο πλαίσιο της μεταβολικής μηχανικής με στόχο τη βελτιωμένη παραγωγή επιθυμητών εκκρινόμενων προϊόντων.

Η αντικειμενική συνάρτηση Z μπορεί να είναι είτε γραμμική είτε μη γραμμική. Η γραμμική χρησιμοποιείται για τη μεγιστοποίηση της βιομάζας ($Z = v_{\text{biomass}}$) ή παραγωγής ATP ($Z = v_{\text{ATP}}$), όπου η ροή (v) μέσω της αντίδρασης παραγωγής συστατικών βιομάζας ή ATP μεγιστοποιείται. Όλες οι μέθοδοι που ανήκουν σε αυτή την κατηγορία υπολογίζουν βέλτιστους φαινότυπους.

Μοναδικό βέλτιστο (Single optima)

Από τη στιγμή που ορίζεται η επιθυμητή λειτουργία του δικτύου η εύρεση της βέλτιστης λειτουργίας ενός GENRE είναι, σε μαθηματικούς όρους, ένα πρόβλημα βελτιστοποίησης με περιορισμούς. Συχνά, τόσο οι περιορισμοί όσο και η αντικειμενική συνάρτηση είναι γραμμικές συναρτήσεις, οπότε και χρησιμοποιείται γραμμικός προγραμματισμός (linear programming-LP). Το αποτέλεσμα της γραμμικής βελτιστοποίησης είναι μία μοναδική κατάσταση (σε μορφή κατανομής ροών) που μεγιστοποιεί το συγκεκριμένο στόχο. Η υπολογισμένη βέλτιστη μορφή του μεταβολικού δικτύου μπορεί να συγκριθεί με πειραματικές μετρήσεις ροών ή

πειραματικά φαινοτυπικά δεδομένα. Ο γραμμικός προγραμματισμός υπολογίζει μία βέλτιστη κατάσταση δικτύου αντιδράσεων. Για τα δίκτυα γονιδιακής κλίμακας, μπορούν να υπάρξουν πολλαπλές καταστάσεις ή κατανομές ροών με την ίδια βέλτιστη λύση της αντικειμενικής συνάρτησης. Δηλαδή ένα μοντέλο γονιδιακής κλίμακας μπορεί να αναπαράξει την ίδια λειτουργία με πολλούς διαφορετικούς τρόπους. Το μαθηματικό αντίστοιχο των ισοδύναμων βέλτιστων καταστάσεων συμπίπτει με τη βιολογική έννοια των «ήσυχων» φαινότυπων (silent phenotypes).



Εικόνα 3.15 Προσδιορισμός βέλτιστων καταστάσεων. Μέσω LP βρίσκουμε μία συγκεκριμένη βέλτιστη λύση. Μέσω MILP βρίσκουμε όλες τις βέλτιστες λύσεις. Μέσω ανάλυσης διακύμανσης ροών βρίσκουμε τα εύρη των τιμών για όλες τις ροές για ένα σετ εναλλακτικού βέλτιστου. Στην εικόνα μόνο η ροή v_1 μεταβάλλεται κατά μήκος του εναλλακτικού βέλτιστου. v_1 : ροή διαμέσου αντίδρασης 1 v_2 : ροή διαμέσου αντίδρασης 2

Εναλλακτικό βέλτιστο (Alternate optima)

Μία ακόμα ιδιότητα των δικτύων γονιδιακής κλίμακας είναι η εναλλακτικές κατανομές ροών που οδηγούν σε ισοδύναμες βέλτιστες καταστάσεις. Ο αριθμός αυτών των βέλτιστων ποικίλλει ανάλογα το μέγεθος του μεταβολικού δικτύου και τις περιβαλλοντικές συνθήκες. Γενικά. Όσο μεγαλύτερο και πιο διασυνδεδεμένο είναι το δίκτυο, τόσο περισσότεροι είναι οι εναλλακτικοί βέλτιστοι φαινότυποι. Έχει αναπτυχθεί ένας αναδρομικός αλγόριθμος μικτού ακεραίου (mixed-integer LP-MILP) για την απαρίθμηση των εναλλακτικών βέλτιστων.

Διακύμανση ροών (Flux variability)

Η ανάλυση διακύμανσης ροών προσδιορίζει το πλήρες εύρος αριθμητικών τιμών για κάθε ροή στο δίκτυο, ενώ παράλληλα ικανοποιούνται οι δοθέντες περιορισμοί και μεγιστοποιείται μία συγκεκριμένη συνάρτηση. Η μέγιστη τιμή της αντικειμενικής συνάρτησης υπολογίζεται πρώτα και ύστερα χρησιμοποιείται ως επιπλέον περιορισμός στο δίκτυο εξασφαλίζοντας ότι μόνο βέλτιστες καταστάσεις δικτύου λαμβάνονται υπόψιν. Ύστερα πραγματοποιούνται πολλαπλές βελτιστοποιήσεις για τον υπολογισμό των μέγιστων και ελάχιστων τιμών των ροών κάθε αντίδρασης.

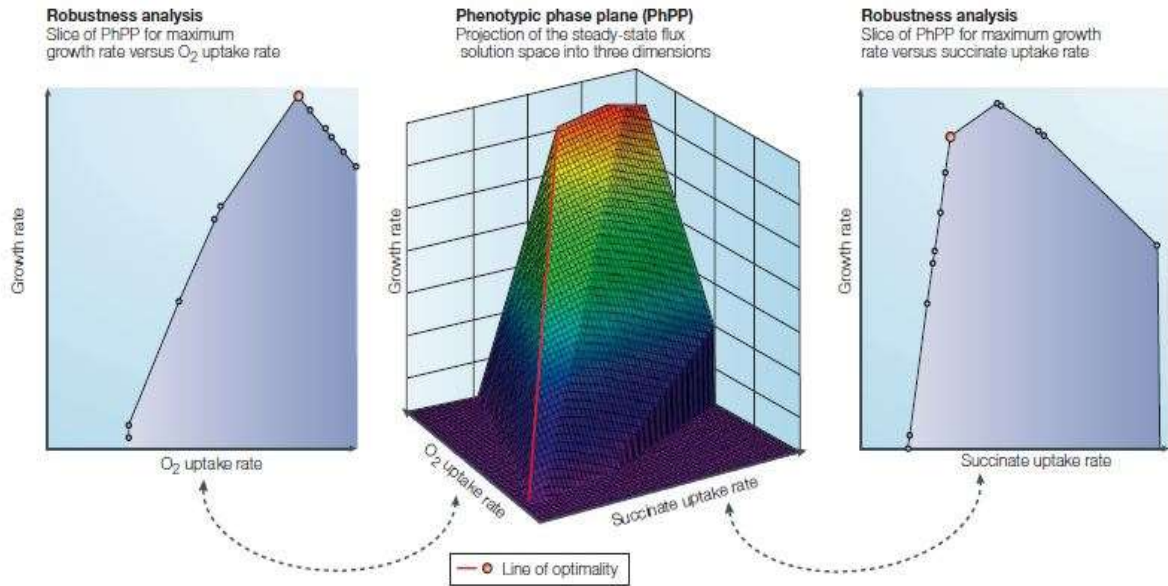
Εύρεση αντικειμενικής συνάρτησης

Ο υπολογισμός της βέλτιστης κατάστασης απαιτεί καθορισμό ενός κυτταρικού σκοπού. Το αντίστροφο πρόβλημα είναι ο υπολογισμός όλων των υποτιθέμενων αντικειμενικών συναρτήσεων στη βάση μετρούμενων εσωκυτταρικών ροών για την εύρεση των αντικειμενικών συναρτήσεων, οι οποίες αν βελτιστοποιηθούν θα έχουν ως αποτέλεσμα μία κατανομή ροών που προσεγγίζει τη μετρούμενη.

Αυτές οι τέσσερις μέθοδοι που αναλύθηκαν επικεντρώνονται στον υπολογισμό βέλτιστων καταστάσεων ενός δικτύου. Ωστόσο, τα βέλτιστα αυτά αποτελέσματα μπορεί να μην παρατηρούνται πειραματικά. Γι' αυτό έχουν αναπτυχθεί άλλες μέθοδοι για τη μελέτη του εύρους κυτταρικών λειτουργιών των μοντέλων γονιδιακής κλίμακας χωρίς να βελτιστοποιούνται αντικειμενικές συναρτήσεις *a priori*.

Μεταβολή μίας παραμέτρου : Ανάλυση Ευρωστίας (Robustness Analysis)

Τα αποτελέσματα έλλειψης-ατελειών ενζύμων στις λειτουργικές καταστάσεις των δικτύων γονιδιακής κλίμακας μπορούν να προσδιοριστούν με ανάλυση ευρωστίας. Η τιμή της ροής περιορίζεται μέσω της επηρεαζόμενης αντίδρασης και η βέλτιστη λύση επανυπολογίζεται μέσα σε αυτό το νέο περιορισμό, ο οποίος αντιπροσωπεύει την ατέλεια ενζύμου. Στην περίπτωση που το ποσό της μείωσης της λειτουργικής ικανότητας του ενζύμου είναι άγνωστο, η ροή μπορεί να αλλάξει διαδοχικά μέσω της αντίδρασης που μας ενδιαφέρει και η αντικειμενική συνάρτηση μπορεί να βελτιστοποιηθεί σε κάθε βήμα.



Εικόνα 3.16 Εξαρτήσεις ροών- Διαγράμματα Ανάλυσης ευρωστίας και επιπέδων φαινοτυπικών φάσεων

Το διάγραμμα που προκύπτει από τη βέλτιστη τιμή έναντι της τιμής ροής της επιθυμητής αντίδρασης δημιουργεί μία καμπύλη γραμμική κατά μέρους. Η κλίση σε κάθε μία γραμμική περιοχή αυτής της καμπύλης αντιπροσωπεύει το μειωμένο κόστος (reduced cost). Το μειωμένο κόστος περιγράφει την αλλαγή στην αντικειμενική συνάρτηση ανά μονάδα αλλαγής στη ροή διαμέσου της αντίδρασης που μας ενδιαφέρει.

Μεταβολή δύο παραμέτρων : Επίπεδα φαινοτυπικών φάσεων (Phenotypic phase planes)

Τα επίπεδα φαινοτυπικών φάσεων αντιπροσωπεύουν μία μέθοδο πραγματοποίησης ανάλυσης ευαισθησίας ως συνάρτηση δύο μεταβλητών. Χρησιμοποιούνται για την απεικόνιση και το χαρακτηρισμό πολλών βέλτιστων καταστάσεων ως συναρτήσεις δύο ροών που μας ενδιαφέρουν. Κάθε λύση μπορεί να χαρακτηριστεί από τις σκιώδεις τιμές (Shadow prices) των δύο ροών που ερευνούνται. Η σκιώδης τιμή είναι ο ρυθμός στον οποίο η αντικειμενική τιμή αλλάζει ως αποτέλεσμα της αύξησης παροχής σε ένα συγκεκριμένο πόρο, για τα μεταβολικά δίκτυα ο πόρος είναι ένας μεταβολίτης. Η δομή της σκιώδους τιμής είναι πεπερασμένη σε αριθμό και διαιρεί το επίπεδο που σχηματίζεται από τα επίπεδα ροών μέσω των δύο αντιδράσεων που μας ενδιαφέρουν σε φάσεις. Οι γραμμές στα επίπεδα φάσεων διαχωρίζουν αυτές τις φάσεις, που αντιστοιχούν στις αλλαγές των σκιωδών τιμών. Με άλλα λόγια, η οριακή

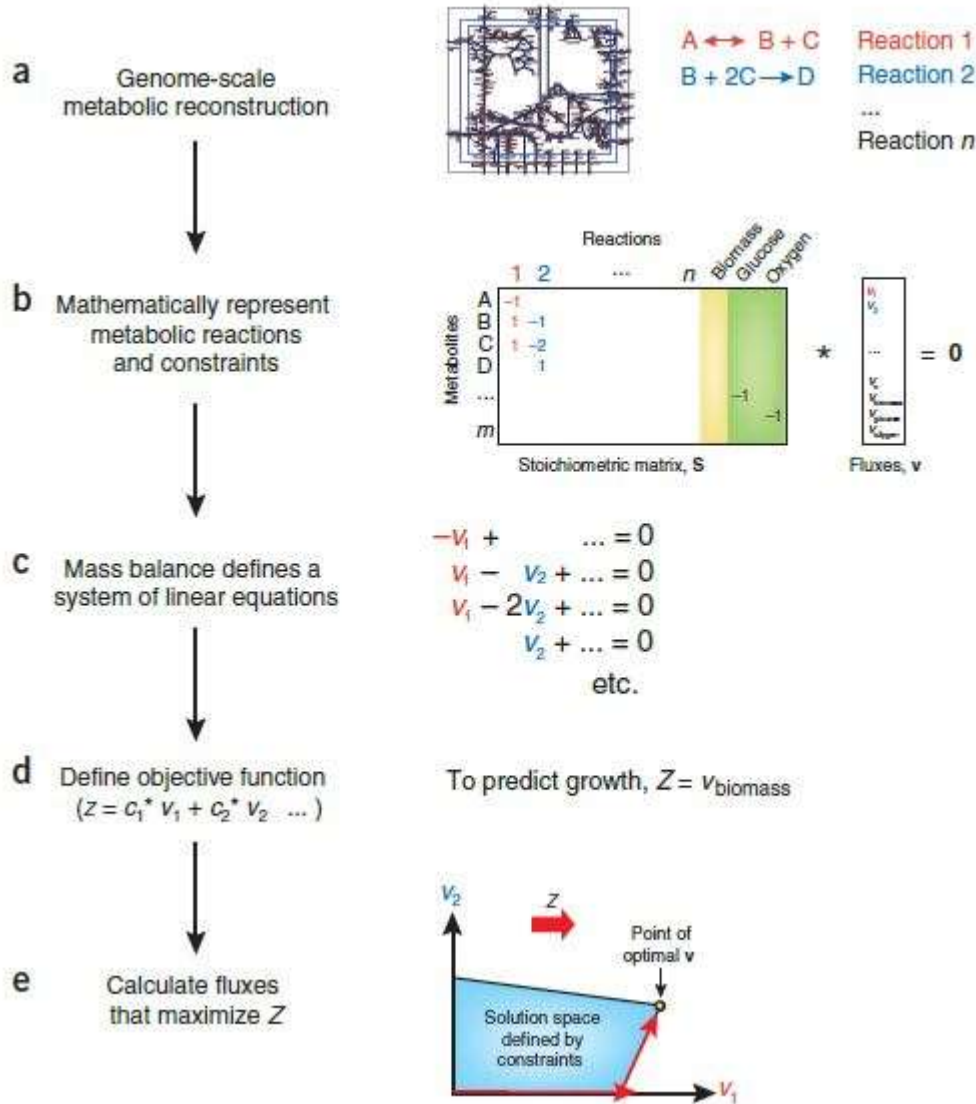
αλλαγή στην τιμή της αντικειμενικής συνάρτησης που συμβαίνει λόγω της οριακής αλλαγής στη διαθεσιμότητα των μεταβολιτών είναι διαφορετική σε κάθε φάση. Οι γραμμές που διαχωρίζουν τις φάσεις μπορούν να έχουν διαφορετικές ονομασίες, όπως η γραμμή βελτιστοποίησης που δείχνει τις καταστάσεις που αντιστοιχούν στη μέγιστη απόδοση βιομάζας ($\text{g DW cell mmol}^{-1}$ πηγή άνθρακα, όπου DW το ξηρό βάρος). Οι φάσεις στα επίπεδα φαινοτυπικών φάσεων μπορούν να κατηγοριοποιηθούν ως : πρώτο,μάταιο- μία αύξηση σε όποια από τις δύο ροές προκαλεί μείωση της αντικειμενικής, δεύτερη, περιορισμός από ένα υπόστρωμα – αύξηση σε μόνο μία από τις ροές αυξάνει την αντικειμενική και τρίτη, περιορισμός από δύο υποστρώματα- αύξηση σε οποιαδήποτε από τις δύο ροές αυξάνει την αντικειμενική. (Price, et al., 2004)

3.2 Ανάλυση Ισορροπίας Ροών (Flux Balance Analysis-FBA)

Η ανάλυση ισορροπίας ροών είναι μια υπολογιστική μέθοδος ανάλυσης υπό περιορισμούς που βασίζεται στο γραμμικό προγραμματισμό. Υπολογίζει τις ροές των αντιδράσεων του μεταβολικού μοντέλου μεγιστοποιώντας ή ελαχιστοποιώντας μία αντικειμενική συνάρτηση που αντικατοπτρίζει τον στόχο που εξυπηρετεί το μοντέλο. Η λύση οριοθετείται από τους βιολογικούς περιορισμούς του προβλήματος.

Τα βήματα που ακολουθούνται για τη λύση ενός τέτοιου προβλήματος είναι τα παρακάτω:

1. Προσδιορισμός του μεταβολικού συστήματος, καταγραφή των βιοχημικών αντιδράσεων, σύνθεση τους σε δίκτυο.
2. Μετατροπή του δικτύου σε μαθηματικό μοντέλο μέσω του πίνακα S.
3. Εφαρμογή περιορισμών που διέπουν το πρόβλημα.
4. Προσδιορισμός αντικειμενικής συνάρτησης.
5. Βελτιστοποίηση



Εικόνα 3.17 Στάδια δημιουργίας ενός προβλήματος ανάλυσης ισοροπίας ροών (FBA)

Για τα πρώτα δύο βήματα έχει γίνει αναφορά στο Κεφάλαιο 2, όπου εξηγήθηκε η κατασκευή του στοιχειομετρικού πίνακα S από ένα μεταβολικό δίκτυο. Τα δύο αυτά βήματα εξασφαλίζουν ένα λειτουργικό μοντέλο.

Στο τρίτο βήμα καθορίζονται οι φυσικοί περιορισμοί που διέπουν το πρόβλημα (βιολογικοί, φυσικοχημικοί, κ.α.) οι οποίοι εκφράζονται μαθηματικά. Στην περίπτωση της FBA, οι περιορισμοί που εκφράζονται ως ισότητες αφορούν τα ισοζύγια μάζας ενώ οι ανισότητες αφορούν όρια στις τιμές των ροών αντιδράσεων και την κατεύθυνση τους. Κατά την εφαρμογή των ισοζυγίων μάζας, θεωρείται σταθερή κατάσταση:

$$S \cdot v = 0 \quad [1]$$

Η εξίσωση [1] αποτελεί τον κύριο περιορισμό της μεθόδου και αναπαριστά ένα περιοριστικό μοντέλο. Από αυτήν προκύπτει μία εξίσωση διατήρησης για κάθε μεταβολίτη και το σύνολο των εξισώσεων αυτών απαρτίζουν το σύστημα προς επίλυση.

Στο τέταρτο, προσδιορίζεται μία γραμμική αντικειμενική συνάρτηση της μορφής:

$$Z = w \times v = \Sigma w_i \times v_i \quad [2]$$

Όπου w είναι το διάνυσμα των συντελεστών της αντικειμενικής συνάρτησης και v , το διάνυσμα των ροών. Οι τιμές των συντελεστών αντικατοπτρίζουν το πόσο συμμετέχει κάθε αντίδραση στην αντικειμενική συνάρτηση.

Τέλος, επιβάλλονται επιπλέον περιορισμοί στις ροές των αντιδράσεων με τη μορφή ανισοτήτων και βελτιστοποιείται το πρόβλημα. Η κανονική μορφή ενός τέτοιου προβλήματος είναι:

$$\begin{array}{ll} \text{maximize} & \mathbf{c}^T \mathbf{v} \\ \text{subject to} & \mathbf{S} \mathbf{v} = \mathbf{0} \\ \text{and} & \text{lowerbound} \leq \mathbf{v} \leq \text{upperbound} \end{array}$$

όπου :

S , γνωστόςστοιχειομετρικός πίνακας

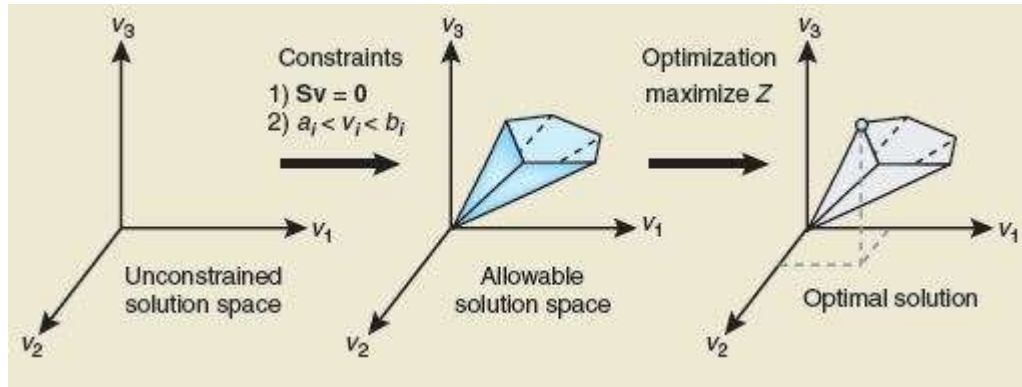
C^T , διάνυσμα πραγματικών αριθμών

v , διάνυσμα των άγνωστων ροών

$Z = C^T \times v$, αντικειμενική συνάρτηση

Είδη λύσεων

Σε ένα πρόβλημα γραμμικού προγραμματισμού συνήθως υπάρχουν άπειρες λύσεις και επιδιώκεται η βέλτιστη μέσω της επιβολής κατάλληλων περιορισμών και μιας αντικειμενικής συνάρτησης. Η εφικτή περιοχή λύσεων βρίσκεται μέσα στον μηδενόχωρο του πίνακα S καθώς αυτός περιγράφει συστήματα σε σταθερή κατάσταση. Μετά την επιβολή περιορισμών, ο μηδενόχωρος συρρικνώνεται και αποκτά τη μορφή ενός κυρτού πολύγωνα όπως φαίνεται στην Εικ.18



Εικόνα 3.18 Επίδραση περιορισμών στο χώρο των λύσεων του συστήματος γραμμικού προγραμματισμού.

Σε ένα πρόβλημα LP υπάρχουν τρεις τύποι δυνατών λύσεων (feasible solutions).

1. Μοναδική Βέλτιστη Λύση (unique solutions).
2. Άπειρες βέλτιστες λύσεις (Degenerate solution).
3. Καμία λύση (unbounded solution)

Βιολογική σημασία λύσεων

1. Μοναδικές Λύσεις. Εμφανίζονται μόνο σε μικρά δίκτυα τα οποία δεν χρειάζονται και δεν έχουν πολλές εναλλακτικές οδούς και έτσι παρουσιάζουν μία βέλτιστη οδό για να επιτύχουν ένα φαινότυπο.

2. Άπειρες Λύσεις. Οι λύσεις αυτές αντιστοιχούν σε ίδιους φαινοτύπους που έχουν αποκτηθεί με διαφορετικό τρόπο. Αποδεικνύουν την ικανότητα ενός οργανισμού να καταλήγει στον ίδιο φαινότυπο από διαφορετικά μονοπάτια πράγμα που συνδέεται με την ομοιόσταση και επιβίωση του κάτω από διαφορετικές συνθήκες. Είναι η συνηθέστερη περίπτωση σε μεγάλης κλίμακας δίκτυα.

3. Καμία Λύση. Δεν αντιστοιχούν σε κάποιο βιολογικά ρεαλιστικό πρόβλημα. Είναι ένδειξη λάθους είτε κατά τη διάρκεια της κατασκευής του δικτύου, είτε κατά την επιβολή των περιορισμών.

(Orth, et al., 2010)

3.3 Βιβλιογραφία

Orth, J. D., Thiele, I. & Palsson, B. Ø., 2010. What is flux balance analysis?. *Nature Biotechnology*, 28(3), pp. 245-248.

Price, N. D., Reed, J. . L. & Palsson, B. Ø., 2004. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nature Reviews Microbiology*, 2(1), pp. 886-897.

ΚΕΦΑΛΑΙΟ 4

ΤΟ ΜΙΚΡΟΦΥΚΟΣ *Nannochloropsis gaditana*

Η παραγωγή βιοκαυσίμων από μικροφύκη κερδίζει συνεχώς έδαφος ως εναλλακτική πηγή ενέργειας λόγω της υψηλής παραγωγικότητας βιομάζας των μικροφυκών καθώς και της μικρής απαίτησής τους σε πόρους. Ωστόσο υπάρχει ακόμα η ανάγκη βελτίωσης της παραγωγικότητας των φυκών για να είναι οικονομικά βιώσιμα τα βιοπροϊόντα από φύκη. Τα μοντέλα μεταβολικών δικτύων των φυκών μπορούν να προσφέρουν ενός είδους «ενόραση» στις στρατηγικές γενετικής τροποποίησης που αποσκοπούν στη βελτίωση στελεχών μικροφυκών. Τα τελευταία χρόνια έχει αναπτυχθεί ένας μεγάλος αριθμός υπολογιστικών εργαλείων που επιτρέπουν ποικιλία αναλύσεων και προβλέψεων και τα οποία βασίζονται σε γενετικούς και θερμοδυναμικούς περιορισμούς ενσωματωμένους στο δίκτυο. Αυτές οι αναλύσεις και προβλέψεις συμβάλλουν στην αναγνώριση εμβιομηχανικών στρατηγικών με αποτέλεσμα βελτιωμένα στελέχη φυκών και βελτιωμένη παραγωγή βιοκαυσίμων. Παρόλο που για ένα ικανό αριθμό γονιδιωμάτων φυκών έχει αναλυθεί η αλληλουχία τους πλήρως, μόνο ελάχιστα μεταβολικά μοντέλα έχουν κατασκευαστεί για αυτά τα είδη.

Η χρησιμότητα των μοντέλων αυτών εκτείνεται σε διάφορες εφαρμογές. Από τη μία βοηθούν στην πλαισίωση υψηλής απόδοσης πειραματικών δεδομένων, όπως για παράδειγμα την ενσωμάτωση δεδομένων έκφρασης γονιδίων με μεταβολικά μονοπάτια κάτω από διαφορετικές συνθήκες ανάπτυξης. Τα μεταβολικά μοντέλα μπορούν ακόμη να αποκαλύψουν στόχους για την μεταβολική μηχανική, οδηγώντας σε αυξημένη παραγωγή του μεταβολίτη-στόχου ή να αυξήσουν κατά προτίμηση τους ρυθμούς αναπνοής. Από την άλλη πλευρά, λόγω της διαθεσιμότητας μεγάλου και ποικίλου μεγέθους βιολογικών δεδομένων, τα μοντέλα μεταβολικών δικτύων παρέχουν ένα πλαίσιο ενσωμάτωσής τους και επιτρέπουν τη διατύπωση και τον έλεγχο μεταγενέστερων υποθέσεων. Επιπλέον η σύγκριση διασταυρούμενων ειδών αντιπροσωπεύει άλλη μία χρησιμότητα τους, διαμέσου της οποίας μπορεί να επιτευχθεί αναγνώριση διαφορεικά ενεργοποιημένων μεταβολικών μονοπατιών. (Koussa J., 2014)

Για την κατασκευή των μοντέλων μικροφυκών χρειάζεται να έχουμε γνώσεις συνθηκών ανάπτυξης και του μεταβολισμού τους.

4.1 Μεταβολισμός φυκών

Τα μικροφύκη αποτελούν δυνητική πηγή για μια μεγάλη ποικιλία προϊόντων, όπως καροτενοειδή, λιπίδια, υδρογόνο, πρωτεΐνες και άμυλο, τα οποία παρουσιάζουν ενδιαφέρον για την παραγωγή διαφόρων προϊόντων, κατά το πλείστον υψηλής αξίας και αφορούν προϊόντα τροφής, κοσμετολογίας, φαρμακευτικής χρήσης, και γενικά βιομηχανική χρήση. (Μάρκου, et al., 2013). Για να παραχθούν αυτές οι ενώσεις απαιτείται καλλιέργεια των μικροφυκών σε συγκεκριμένες συνθήκες.

Τα μικροφύκη είναι μονοκύτταροι ευκαρυωτικοί οργανισμοί που μπορούν να χρησιμοποιήσουν το φως ως πηγή ενέργειας μέσω της φωτοσύνθεσης. Στην πιο ευρεία έννοια περιλαμβάνουν και τα κυανοβακτήρια (μπλε-πράσινα άλγη). (Beardall & Raven, 2012)

Είδη μεταβολισμού

Οι μεταβολικοί τύποι με τους οποίους μπορούν να αναπτυχθούν ορισμένα μικροφύκη είναι: 1) Φωτοαυτότροφος: στον τύπο αυτό τα μικροφύκη αναπτύσσονται χρησιμοποιώντας ως πηγή ενέργειας τον ήλιο και ως πηγή άνθρακα το CO₂ ή άλλες ανόργανες μορφές, 2) Ετερότροφος: Στον τύπο αυτό τα μικροφύκη χρησιμοποιούν ως πηγή ενέργειας και πηγή άνθρακα διάφορες οργανικές ουσίες όπως γλυκόζη, οξικό οξύ, γλυκερίνη κλπ, 3) Μιξότροφος: τα μικροφύκη στον τύπο αυτό είναι σε θέση να αυξάνονται/πολλαπλασιάζονται είτε αξιοποιώντας τον ήλιο ως πηγή ενέργειας είτε αξιοποιώντας διάφορες οργανικές ουσίες και ως πηγή άνθρακα είτε το CO₂ είτε τις οργανικές ουσίες και 4) Φωτοετερότροφος: στον τύπο αυτό τα μικροφύκη χρησιμοποιούν ως πηγή ενέργειας τον ήλιο και σαν πηγή άνθρακα τις οργανικές ενώσεις. Ο ετερότροφος και μιξότροφος μεταβολισμός μπορούν να αποτελέσουν ιδανική καλλιεργητική τεχνική για την παραγωγή μικροφυκών, γιατί έτσι αντιμετωπίζεται το πρόβλημα της αξιοποίησης του φωτός σε καλλιέργειες με μεγάλη πυκνότητα ή σε καλλιέργειες με φωτοπαρεμπόδιση (κατά τους καλοκαιρινούς μήνες). Στις ετερότροφες ή/και μιξότροφες καλλιέργειες η προσθήκη οργανικού άνθρακα στο υπόστρωμα καλλιέργειας δίνει μεγαλύτερη συγκέντρωση βιομάζας σε σχέση με τις φωτοαυτότροφες καλλιέργειες. Γενικά, ο μιξότροφος μεταβολισμός έχει

μεγαλύτερους ρυθμούς ανάπτυξης και μεγαλύτερη τελική παραγωγή σε βιομάζα σε σχέση με τους άλλους τρεις μεταβολισμούς (Μάρκου, et al., 2013)

Πολλές από τις μεταβολικές διεργασίες που συμβαίνουν στα άλγη είναι κοινές με άλλων ζώντων οργανισμών, αλλά υπάρχουν μηχανισμοί όπως η συγκέντρωση φωτός, η απόκτηση άνθρακα, αφομοίωση αζώτου και θείου καθώς και σχηματισμός μοναδικών δευτερογενών μεταβολιτών που είναι χαρακτηριστικοί για τα άλγη.

Τα άλγη είναι ικανά να φωτοσυνθέτουν (oxygen-evolving photosynthesis) αλλά έχουν έλλειψη αναπαραγωγικών και αναπτυξιακών χαρακτηριστικών που σχετίζονται με τα ανώτερα φυτά και εμβρυόφυτα. Τα ευκαρυωτικά άλγη, όπως είναι και το φύκος *Nannochloropsis*, απέκτησαν τη φωτοσυνθετική τους μηχανή (χλωροπλάστης) από την κατάποση, διατήρηση και γενετική ενσωμάτωση κυανοβακτηρίων ή προϋπάρχοντων ευκαρυωτικών οργανισμών που εξελίχθηκαν με οξυγόνο από φαγοτροφικούς, μη φωτοσυνθετικούς ευκαρυώτες. Επομένως, ως φωτοσυνθετικοί οργανισμοί, τα ευκαρυωτικά άλγη είναι πολυφυλετικά, παρόλο που η φωτοσυνθετική τους μηχανή είναι μονοφυλετική. Αυτό σημαίνει ότι μπορεί να υπάρχει ποικιλομορφία μεταβολικών διεργασιών ως αποτέλεσμα της εξελικτικής τους ιστορίας. Μέχρι σήμερα έχει υπάρξει ραγδαία εξέλιξη στην αλληλούχηση γονιδιωμάτων διάφορων φυκών. Η αποκωδικοποίηση των πληροφοριών από τα γονιδιώματα αυτά δίνει σημαντική γνώση για τη βιοχημεία και το μεταβολισμό των φυκών.

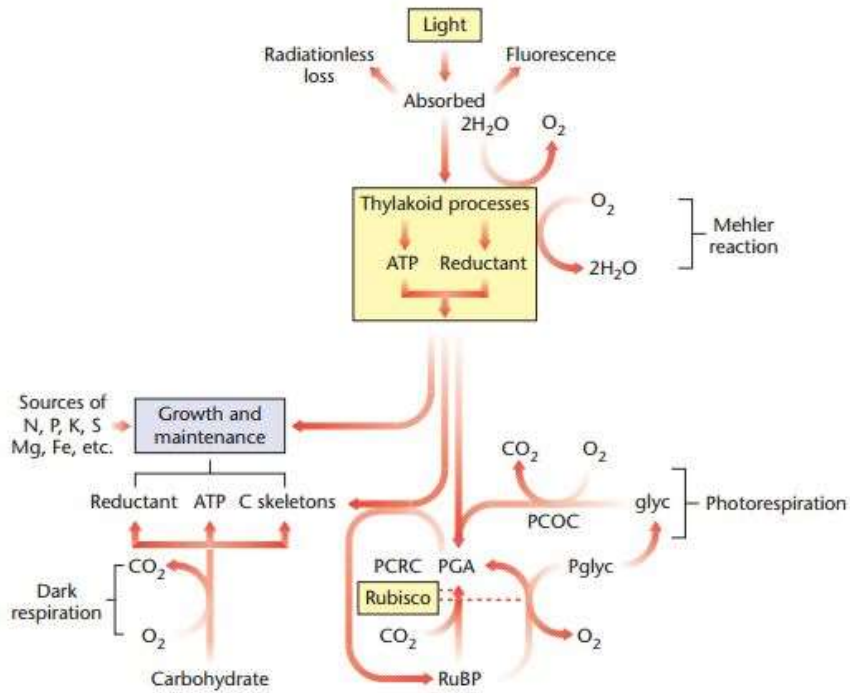
Ένα βασικό χαρακτηριστικό που διαχωρίζει τα άλγη από κατώτερους ευκαρυώτες είναι η οξυγονική φωτοσύνθεση. Σημαντικά στοιχεία της φωτοσυνθετικής διαδικασίας είναι η αντιδράσεις που καταλύονται από δύο κέντρα σύμπλοκων αντιδράσεων, το σύμπλοκο κυτοχρώματος b_6-f , διάφοροι παράγοντες M_f οξειδοαναγωγής, τη συνθετάση τριφωσφορικής αδενοσίνης (ATPase) και ένζυμα της φωτοσυνθετικής αναγωγής άνθρακα.

Οι φωτοσυνθετικές αντιδράσεις παράγουν σάκχαρα τα οποία χρησιμοποιούνται από της χημο-οργανοτροφικές αντιδράσεις των κυανοβακτηρίων και ευκαρυωτών που έχουν αποκτήσει χλωροπλάστες. Ο βασικός χημο-οργανοτροφικός μεταβολισμός των άλγων είναι γενικά αρκετά όμοιος με αυτόν των ευ-βακτηρίων και ευκαρυώτων, αν και υπάρχουν εξαιρέσεις που παρουσιάζονται παρακάτω.

Ο φωτολιθοτροφικός μεταβολισμός (φωτολιθότροφοι οργανισμοί αποκτούν ενέργεια από το φως και συνεπώς χρησιμοποιούν ανόργανους δέκτες ηλεκτρονίων μόνο για να τροφοδοτήσουν βιοσυνθετικές αντιδράσεις) των φυκών συμβαίνει στα υποχρεωτικά φωτολιθοτροφικά άλγη, τα οποία δεν μπορούν να αναπτυχθούν στο σκοτάδι σε κανένα οργανικό υπόστρωμα που έχει ελεγχθεί μέχρι τώρα.

Η ικανότητα ανάπτυξης με φωτόνια, διοξείδιο του άνθρακα, άζωτο ως αέριο (N_2), αμμώνιο (NH_4^+) ή νιτρικό (NO_3^-), δισόξινο φωσφορικό ($H_2PO_4^-$), θειϊκό άλας (SO_4^{2-}) και άλλα απαραίτητα στοιχεία που προμηθεύονται σε ανόργανη μορφή εκτείνεται στ άλγη που δεν είναι υποχρεωτικά φωτολιθοτροφικά, οπότε μπορούν να αναπτυχθούν στο φως, και λιγότερο συχνά στο σκοτάδι με οργανικό άνθρακα ως βασική πηγή άνθρακα. Η απώλεια λειτουργικότητας ενός ή περισσότερων γονιδίων για τη φωτοσύνθεση οδηγεί σε υποχρεωτικά χημοργανοτροφικό μεταβολισμό, ο οποίος μεταξύ ευκαρυωτικών οργανισμών περιλαμβάνει φαγοτροφία ή παρασιτισμό. Το εύρος των ενώσεων οργανικού άνθρακα που μπορούν να χρησιμοποιηθούν από χημοργανοτροφικά ή φωτοργανοτροφικά άλγη ποικίλλει ανάλογα με την ταξονομία τους αλλά συνήθως περιλαμβάνει γλυκόζη, οξικό άλας και κάποια αμινοξέα. Όσον αφορά την παροχή αζώτου η χρήση N_2 ως πηγή αζώτου είναι εφικτή σε συγκεκριμένα κυανοβακτήρια. Στα ευκαρυωτικά άλγη η δέσμευση αζώτου σχετίζεται με βακτηριακούς δεσμευτές.

Άλλες αποκλίσεις από την αυστηρή φωτολιθοτροφία σχετίζονται με βλάβες στις μεταβολικές διεργασίες που κάνουν κάποια άλγη εξαρτώμενα από εξωγενή τροφοδοσία συγκεκριμένων ουσιών ανάπτυξης ή βιταμινών, κυρίως βιοτίνη, θιαμίνη και βιταμίνη B_{12} .



Εικόνα 4.1 Σχεδιάγραμμα μονοπατιών ενέργειας, άνθρακα, οξυγόνου στη φωτοσύνθεση, φωτοαναπνοή, και ανάπτυξη ενός φύκου. Δεν παρουσιάζονται στοιχειομετρίες. (Συντομογραφίες: Glyc, υδροξυαιθανικό οξύ; PCOC, κύκλος δέσμευσης άνθρακα με φωτοαναπνοή; PCRC, κύκλος φωτοσυνθετικής αναγωγής άνθρακα; PGA, 3-φωσφογλυκερικό οξύ; Pglyc, φωσφοδροξυαιθανικό οξύ; RuBP, ribulose biphosphate; Rubisco, ribulose biphosphate carboxylaseoxygenase

Λειτουργίες που συμβαίνουν μόνο στα άλγη

Σημαντικά μάκρο- και μικρομόρια, τα οποία παράγονται ειδικά από άλγη, κυμαίνονται από πρωτεϊνικά σύμπλοκα χρωστικών μέχρι οξειδοαναγωγικούς καταλύτες της φωτοσύνθεσης, ένζυμα και μεταφορείς που σχετίζονται με δέσμευση ανόργανου άνθρακα και ένζυμα του μεταβολισμού υδροαιθανικού οξέος. (Beardall & Raven, 2012)

4.2 Καλλιέργεια μικροφυκών

Η καλλιέργεια των μικροφυκών μπορεί να πραγματοποιηθεί σε ανοικτούς ή κλειστούς τύπους καλλιέργειών. Στις ανοικτού τύπου καλλιέργειες, τα μικροφύκη καλλιεργούνται σε τεχνητούς ή φυσικούς περιέκτες, όπως δεξαμενές, νερόλακκοι, λίμνες κλπ. Στις κλειστού τύπου καλλιέργειες η καλλιέργεια γίνεται εντός διαφόρων σχημάτων κλειστών περιεκτών που ονομάζονται φωτοβιοαντιδραστήρες (photobioreactors). Οι φωτοβιοαντιδραστήρες αποτρέπουν την άμεση επαφή της καλλιέργειας με τις εξωτερικές συνθήκες και έτσι την προστατεύουν από τους διάφορους περιβαλλοντικούς κινδύνους.

Ανοικτού τύπου

Οι ανοικτού τύπου καλλιέργειες έχουν το πλεονέκτημα ότι έχουν σχετικά μικρότερο κόστος εγκατάστασης και λειτουργίας από τις κλειστού τύπου καλλιέργειες και έχουν σχετικά καλή δυνατότητα μαζικής καλλιέργειας μικροφυκών. Στις ανοικτού τύπου καλλιέργειες όμως τα μικροφύκη είναι εκτεθειμένα στους περιβαλλοντικούς παράγοντες και συνήθως επιμολύνονται από διάφορα άλλα είδη μικροφυκών ή βακτηρίων ή άλλων μικροοργανισμών. Οι επιμολύνσεις με άλλους μικροοργανισμούς αλλοιώνει την σύσταση των καλλιέργειών και δυσχεραίνει την εφαρμογή της μονοκαλλιέργειας. Έντονες επιμολύνσεις μπορεί να έχουν ως αποτέλεσμα την μικρή παραγωγή ή και σε ορισμένες περιπτώσεις την κατάρρευση της καλλιέργειας είτε λόγω της επικράτησης άλλων μικροοργανισμών είτε λόγω της εμφάνισης παρασιτικών ασθενειών. Οι ανοικτού τύπου καλλιέργειες πάσχουν επιπλέον λόγω της εξάτμισης νερού που έχει ως αποτέλεσμα την απώλεια σημαντικών ποσοτήτων νερού και την αύξηση της συγκέντρωσης αλάτων που από ένα επίπεδο και πάνω επιδρούν ανασταλτικά στην αύξηση των μικροφυκών. Το γεγονός αυτό οδηγεί στην ανάγκη αναπλήρωσης νερού άρα και χρήση μεγάλων ποσοτήτων του.

Επίσης η παραγωγή ανά μονάδα επιφάνειας είναι μικρότερη σε σχέση με τις κλειστού τύπου καλλιέργειες με αποτέλεσμα η συγκέντρωση της βιομάζας να είναι σχετικά μικρή ανεβάζοντας πολύ το κόστος συγκομιδής. Επίσης σε ορισμένες τοποθεσίες, οι περίοδοι καλλιέργειας είναι περιορισμένες, λόγω έντονων εποχιακών περιβαλλοντικών διακυμάνσεων ή χαμηλών θερμοκρασιών και ηλιοφάνειας.

Κλειστού τύπου

Στις καλλιέργειες κλειστού τύπου, οι συνθήκες είναι περισσότερο ή λιγότερο ελεγχόμενες, αναλόγως αν πρόκειται για καλλιέργειες στεγασμένες ή υπαίθριες. Στις υπαίθριες καλλιέργειες οι συνθήκες δεν ελέγχονται πλήρως (ηλιακή ένταση και θερμοκρασία). Αντίθετα στις στεγασμένες καλλιέργειες οι συνθήκες με τον ένα ή άλλο τρόπο είναι πιο ελεγχόμενες. Οι κλειστού τύπου καλλιέργειες πραγματοποιούνται σε κλειστούς περιέκτες, τους φωτοβιοαντιδραστήρες (photobioreactors).

Τα σημαντικότερα πλεονεκτήματα των φωτοβιοαντιδραστήρων είναι: 1) καλύτερος έλεγχος των συνθηκών καλλιέργειας, 2) μεγαλύτερος λόγος επιφάνειας/όγκου, 3) σε ορισμένους τύπους φωτοβιοαντιδραστήρων καλύτερος έλεγχος της παροχέτευσης και της κίνησης των αερίων (πχ CO₂), 4) περιορισμός της εξάτμισης νερού από το υπόστρωμα καλλιέργειας, 5) καλύτερη θερμική κατανομή στο υπόστρωμα καλλιέργειας, 6) σχετικά εύκολη εγκατάσταση σε ανοιχτό χώρο, 7) προστασία από επιμολύνσεις από διάφορους μικροοργανισμούς 8) σχετικά υψηλές αποδόσεις σε σχέση με τις ανοικτές δεξαμενές και 9) η δυνατότητα καλλιέργειας σε τύπο μιξοτροφίας ή ετεροτροφίας χωρίς ιδιαίτερα προβλήματα επιμόλυνσης. (Μάρκου, et al., 2013)

4.3 Το θαλάσσιο μικροφύκος *Nannochloropsis gaditana*

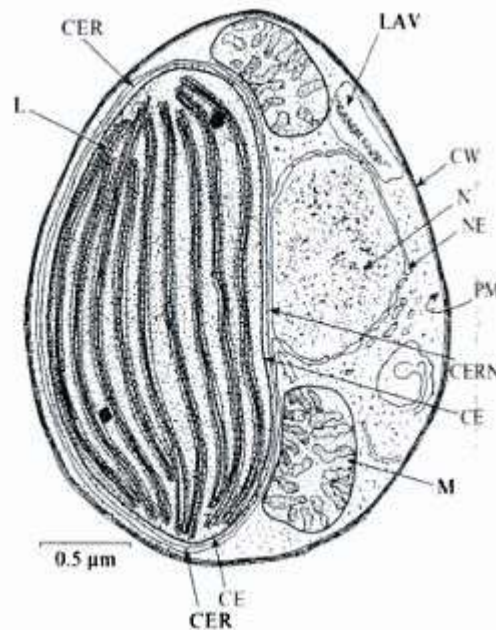
4.3.1 Ταξονομία και βασικά χαρακτηριστικά

Το *Nannochloropsis* είναι ένα γένος φυκών που αποτελείται από έξι είδη, τα οποία είναι τα εξής αναγνωρισμένα:

- *N. gaditana*
- *N. granulata*
- *N. limnetica*
- *N. oceanica*
- *N. oculata*
- *N. salina*

(D.J.Hibberd, 1981)

Το φύκος αυτό ανήκει στην τάξη των Ευστιγματόφυτων (Eustigmatophyceae), και έχει διαχωριστεί από τα Ξανθόφυτα (Xanthophyceae). Τα ευστιγματόφυτα είναι μικρά μικροφύκη, μονοκύτταρα τα οποία είναι πλανγκτονικά ή μεγαλώνουν προσκολλημένα σε διάφορα υποστρώματα. Είναι μοναδικά λόγω του γεγονότος ότι δεν παράγουν χλωροφυλλικές χρωστικές ουσίες πέραν της χλωροφύλλης α. (Fischer, et al., 1998) (Sukenik, et al., 1989). Το φύκος *Nannochloropsis* έχει σφαιρικό, ελαφρώς ωοειδές σχήμα και περιέχει ένα χλωροπλάστη ανά κύτταρο. Η βασική χρωστική που παράγει είναι η βιολαξανθίνη. (Hoek, et al., 1995)



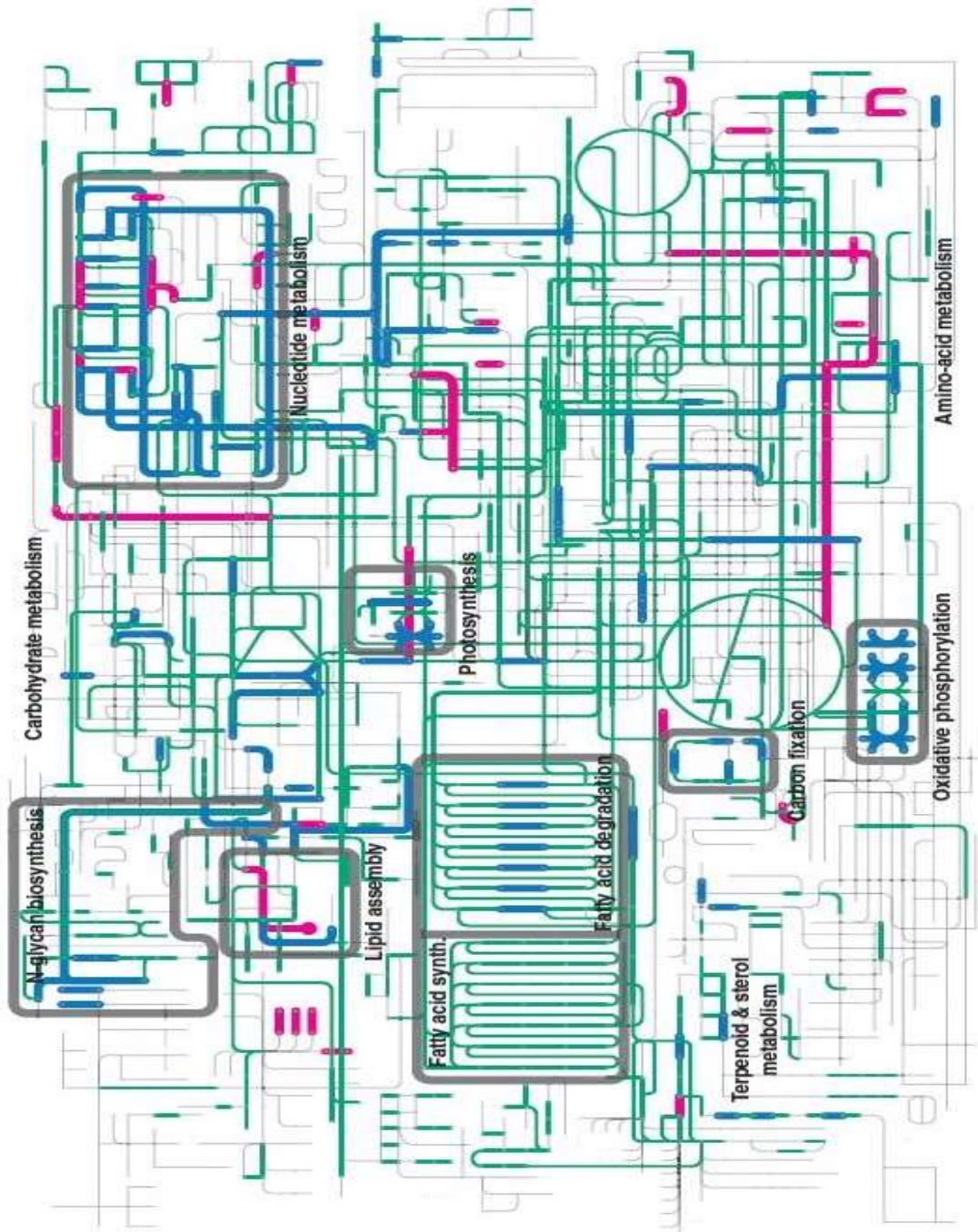
Εικόνα 4.2 Αντιπροσωπευτική εικόνα ενός κυττάρου του *Nannochloropsis*. (Hoek, et al., 1995)

Τα είδη του γένους *Nannochloropsis* χαρακτηρίζονται από μοναδικά βιοχημικά και δομικά στοιχεία όπως η έλλειψη χλωροφύλλης b και c, η σύνθεση των κυτταρικών χρωστικών και το υψηλό περιεχόμενο σε εικοσαπενταενοϊκό οξύ. Αυτή η υψηλή παραγωγικότητα του *Nannochloropsis* σε λιπίδια τον καθιστά ένα πολλά υποσχόμενο οργανισμό για εφαρμογές βιοκαυσίμων. Πιο συγκεκριμένα το είδος *N. gaditana* έχει υψηλούς ρυθμούς παραγωγής φωτοαυτοτροφικής βιομάζας και λιπιδίων, μπορεί να αναπτύσσεται σε μεγάλες πυκνότητες και παρουσιάζει μεγάλη ανοχή σε μεγάλο εύρος συνθηκών σχετικά με το pH, τη θερμοκρασία και την αλατότητα. Αυτό τον καθιστά καλό υποψήφιο για πρότυπο οργανισμό στην παραγωγή βιοκαυσίμων από άλγη. Στην Εικ.3 απεικονίζεται ο χάρτης των μεταβολικών μονοπατιών του *N. Gaditana* όπως δίνεται από την βάση KEGG. (Radakovits, et al., 2012)

4.3.2 Μεταβολισμός

Ο μεταβολισμός του *N. gaditana* αποτελείται από τα εξής βασικά μονοπάτια, τα οποία επιχειρούμε να μοντελοποιήσουμε στη συνέχεια:

- Μεταβολισμός αμινοξέων.
- Μεταβολισμός υδατανθράκων-κύκλος κιτρικού οξέος, μεταβολισμός φρουκτόζης, μαννόζης, γαλακτόζης, αμύλου και σουκρόζης κ.α..
- Οξειδωτική φωσφορύλιωση και φωτοσύνθεση μαζί με την φωτοαναπνοή, δέσμευση άνθρακα, το μεταβολισμό θείου, αζώτου και μεθανίου για τον ενεργειακό μεταβολισμό.
- Μεταβολισμός λιπιδίων- σύνθεση-αποδόμηση λιπαρών οξέων και πολυακόρεστων λιπαρών οξέων καθώς και σύνθεση φωσφο-, γλυκεροφωσφολιπιδίων και βιοσύνθεση τερπενοειδών.
- Μεταβολισμό νουκλεοτιδίων- πυριμιδίνης και πουρίνης.
- Μεταβολισμό τερπενοειδών και πολυκετιδίων-περιλαμβάνει τη σύνθεση καροτενοειδών.
- Μεταβολισμός βιταμινών και συμπαραγόντων-θιαμίνη, ριβοφλαβίνη, βιοτίνη καθώς και μεταβολισμός πορφυρινών και χλωροφύλλης.
- Βιοσύνθεση δευτερογενών μεταβολιτών.



Εικόνα 4.3 Χάρτης μεταβολικών μονοπατιών για το *N. gaditana*. Οι ανοιχτές γκρι γραμμές υποδηλώνουν μονοπάτια KEGG που δεν κωδικοποιούνται από το γονιδίωμα του οργανισμού. Τα μονοπάτια σε μπλε, ροζ και γκρι χρώμα είναι παρόντα στον οργανισμό. Τα γονίδια που είναι αύξο- ή μειορυθμιζόμενα κατά την έλλειψη αζώτου χρωματίζονται με ροζ και μπλε αντίστοιχα. (Radakovits, et al., 2012)

4.3.3 Μεταβολικό προφίλ του *Nannochloropsis* υπό συνθήκες έλλειψης αζώτου

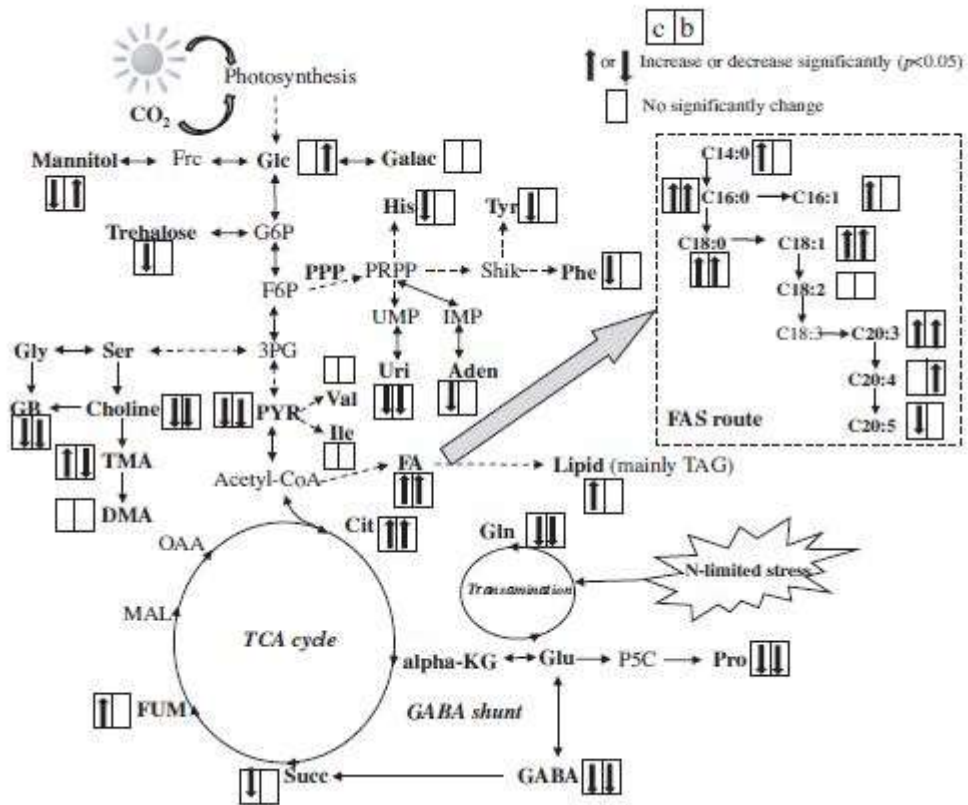
Το ευκαρυωτικό θαλάσσιο μικροφύκος *Nannochloropsis* όπως έχουμε αναφέρει έχει μελετηθεί ως ισχυρός υποψήφιος για την παραγωγή βιοκαυσίμων και σύμφωνα με έρευνες το λιπιδικό του περιεχόμενο είναι πάνω από 30%, καθιστώντας αυτό το είδος τον καλύτερο παραγωγό λιπιδίων ανάμεσα σε 30 θαλάσσια μικροφύκη από πλευρά παραγωγικότητας λιπιδίων. Λόγω του υψηλού περιεχομένου του σε εικοσαπεντανοϊκό οξύ (EPA) έχει προταθεί ως εναλλακτική πηγή εμπορικού EPA. Προς το παρόν καλλιεργείται επιτυχώς σε εσωτερικά και εξωτερικά συστήματα για παραγωγή βιοντίζελ ή βιομάζας. Ωστόσο, η συσσώρευση ελαίων στα κύτταρα φυκών, η οποία συχνά επιτυγχάνεται μέσω έλλειψης αζώτου, συνοδεύεται από χαμηλή παραγωγικότητα βιομάζας, κάτι που είναι αντιφατικό ως προς εξασφάλιση υψηλής ποσότητας βιομάζας και λιπιδίων σε ένα βήμα.

Οι στρατηγικές περιβαλλοντικού στρες που έχουν αναπτυχθεί επιτυχώς για τον χειρισμό του μεταβολισμού του *Nannochloropsis* προς μεγαλύτερη παραγωγή λιπιδίων (πρωτίστως τριγλυκεριδίων) είναι ο έλεγχος της συγκέντρωσης αερισμού με CO₂, ένταση φωτός, αλατότητα, θερμοκρασία και συγκέντρωση αζώτου.

Με βάση την αλληλουχία γονιδιώματος του *Nannochloropsis* που είναι πλέον διαθέσιμη, το μονοπάτι και η ρύθμιση της βιογένεσης λιπιδίων έχουν περιγραφεί λεπτομερώς και ο χάρτης του μεταβολικού μονοπατιού είναι πλέον πιο ξεκάθαρα προσδιορισμένος.

Η πιο συνηθισμένη μέθοδος συσσώρευσης λιπιδίων στα άλγη είναι η έλλειψη αζώτου. Πειράματα σε διάφορες συνθήκες έλλειψης αζώτου σε καλλιέργειες παρτίδας αποδεικνύουν ότι τα λιπίδια συσσωρεύονται σημαντικά συναρτήσει του χρόνου, αλλά οι πρωτεΐνες και οι υδατάνθρακες εμφανίζουν μικρή διακύμανση με μείωση και αύξηση αντίστοιχα. Η αύξηση των λιπιδίων οφείλεται κυρίως στην αύξηση των ουδέτερων λιπιδίων ενώ τα φωσφωλιπίδια στο κύτταρο, ή κάποια ποσότητά τους, καταναλώνονται. Είναι γνωστό ότι τα τριγλυκερίδια, τα οποία είναι βασικά συστατικά των ουδέτερων λιπιδίων βιοσυντίθενται μέσω δύο μονοπατιών. Το άμεσο μονοπάτι γλυκερόλης που περιλαμβάνει την acyl-CoA εξαρτώμενη ακυλτρανσφεράση διακυλγλυκερίνης (DAGAT). Σε ένα άλλο acyl-CoA εξαρτώμενο μονοπάτι, τα πολικά λιπίδια μετατρέπονται σε τριγλυκερίδια με μία τρανς-

εστεροποίησης που καταλύεται από το ένζυμο PDAT. Η σχετική συνεισφορά αυτού του ενζύμου έναντι του DAGAT στην σύνθεση τριγλυκεριδίων εξαρτάται από τη φάση ανάπτυξης του κυττάρου. Και τα δύο αυτά ένζυμα είναι παρόντα στον οργανισμό του *Nannochloropsis*. Παρατηρήθηκε ότι το επίπεδο των πολικών λιπιδίων μειώθηκε με έλλειψη αζώτου, υποδεικνύοντας την ύπαρξη του ενζύμου PDAT και του αντίστοιχου μονοπατιού. Από τα αποτελέσματα των πειραμάτων συμπεραίνεται ότι για μεγαλύτερη απόδοση ουδέτερων λιπιδίων για παραγωγή βιοντίζελ η καλύτερη συνθήκη είναι παροχή χαμηλής συγκέντρωσης αζώτου, αλλά για τη μεγαλύτερη απόδοση EPA, η υψηλή συγκέντρωση αζώτου είναι η βέλτιστη συνθήκη. (Yan, et al., 2013)



Εικόνα 4.4 Απόκριση του *Nannochloropsis* σε διάφορες συγκεντρώσεις αζώτου (Yan, et al., 2013)

4.4 Βιβλιογραφία

- Beardall, J. & Raven, J. A., 2012. Algal Metabolism. *Encyclopedia of Life Sciences*.
- D.J.Hibberd, 1981. Notes on the taxonomy and nomenclature of the algal classes Eustigmatophyceae and Tribophyceae (synonym Xanthophyceae). *Issue Botanical Journal of the Linnean Society*, 82(2), pp. 93-119.
- Fischer, T., Berner, T., Iluz, D. & Dubinsky, Z., 1998. The kinetics of photoacclimation response of *Nannochloropsis* sp (Eustigmatophyceae): A study of changes in ultrastructure and PSU density. *Journal of Phycology*, Τόμος 34, pp. 818-824.
- Hoek, C., Mann, D. & Jahns, H. M., 1995. Heterokontophyta: Class Eystigmatophyceae. Στο: G. Thieme, επιμ. *Algae: An Introduction to Phycology*. Cambridge: Cambridge University Press.
- Koussa J., C. A.-A. K., 2014. Computational Approaches for Microalgal Biofuel Optimization: A Review. *BioMed Research International*, Τόμος 2014, pp. 1-3.
- Radakovits, R. και συν., 2012. Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*. *Nature Communications*, 3(1).
- Sukenik, A., Carmeli, Y. & Berner, T., 1989. Regulation of fatty acid composition by irradiance level in the eustigmatophyte *Nannochloropsis* sp.. *Journal of Phycology*, Τόμος 25, pp. 686-692.
- Yan, X. και συν., 2013. Metabolic profiles of *Nannochloropsis oceanica* IMET1 under nitrogen-deficiency stress. *Bioresource Technology*, 130(1), pp. 731-738.
- Μάρκου, Γιώργος, Τ., Ιωάννης, Ν. & Ηλίας, 2013. Μικροφύκη: Καλλιέργεια και βιομηχανικές εφαρμογές. *e-Journal of Science & Technology*, 8(1), pp. 9-29.

ΚΕΦΑΛΑΙΟ 5

ΚΑΤΑΣΚΕΥΗ ΜΕΤΑΒΟΛΙΚΟΥ ΜΟΝΤΕΛΟΥ ΤΟΥ *Nannochloropsis gaditana*

5.1 Εργαλεία

- ❖ Υπολογιστικό πακέτο Matlab R2012a.
- ❖ Υπολογιστικό εργαλείο COBRA Toolbox που ενσωματώνεται στο Matlab.
- ❖ Υπολογιστικό εργαλείο SBML Toolbox που ενσωματώνεται στο Matlab και επιτρέπει την ανάγνωση μοντέλων σε μορφή SBML.
- ❖ Εργαλείο γραμμικού προγραμματισμού glpk solver, ενσωματώνεται στο Matlab, είναι συμβατό με το COBRA Toolbox και διατίθεται δωρεάν.
- ❖ Excel, για τη συλλογή πληροφοριών κατασκευής προσχεδίου.

5.2 Διαδικασία κατασκευής μοντέλου

Η βάση για την κατασκευή του μεταβολικού μοντέλου του *Nannochloropsis* ήταν το μοντέλο γονιδιακής κλίμακας AlgaGEM. Το AlgaGEM καλύπτει το μεταβολισμό ενός διαμερισματοποιημένου (σε κυτταρόπλασμα, μιτοχόνδριο, πλασμίδιο κλπ) κυττάρου φύκους και βασίστηκε στο γονιδίωμα του *Chlamydomonas reinhardtii*. (Gomes de Oliveira Dal'Molin, et al., 2011).

Παρόλο που η βάση ήταν ένα ήδη υπάρχον μοντέλο, για την κατασκευή του νέου μοντέλου ακολουθήθηκε το πρωτόκολλο κατασκευής μοντέλων γονιδιακής κλίμακας και τα στάδια που αυτό περιλαμβάνει όπως αναφέρθηκαν και σε παραπάνω κεφάλαιο. Παρακάτω περιγράφονται πιο αναλυτικά τα βήματα για την ολοκλήρωση του κάθε σταδίου. Τα βήματα που δίνονται με αστεράκι (*) δεν περιλαμβάνονται στο πρωτόκολλο αλλά ήταν απαραίτητα για την κατασκευή του μοντέλου.

1^ο ΣΤΑΔΙΟ : ΔΗΜΙΟΥΡΓΙΑ ΠΡΟΧΕΔΙΟΥ (Draft Reconstruction)**1| Συλλογή γονιδιακών δεδομένων και αναγνώριση υποψήφιων μεταβολικών λειτουργιών.**

Το πρώτο βήμα είναι η ανάκτηση της λίστας κωδικών ενζύμων (EC numbers), των αντίστοιχων γονιδίων, καταλυτικής δράσης και υποκυτταρικής περιοχής για το *Nannochloropsis gaditana* από τη βάση δεδομένων Uniprot. Στόχος είναι η σύγκριση των ενζύμων του *Nannochloropsis* με αυτά που χρησιμοποιήθηκαν στο AlgaGEM για να εξαχθεί η λίστα εκείνων που πρέπει να συμπληρωθούν για την κατασκευή του μοντέλου. Τα δεδομένα αυτά λήφθηκαν από τη βάση δεδομένων σε μορφή excel.

Στη συνέχεια εξήχθη το μοντέλο του AlgaGEM σε μορφή excel, μέσω του Matlab έτσι ώστε να γίνει εξαγωγή της λίστας με τους κωδικούς ενζύμων του μοντέλου. Δυστυχώς επειδή η συμπλήρωση των κωδικών αυτών είναι προαιρετική δεν περιλαμβάνονταν στο αρχείο, αλλά βρέθηκε αρχείο για το μοντέλο αυτό που αντιστοιχούσε τις αντιδράσεις του με κωδικούς ενζύμων. Λόγω του μεγάλου όγκου δεδομένων η σύγκριση αυτοματοποιήθηκε μέσω κώδικα στο Matlab ο οποίος ήταν ο εξής:

```

% Εύρεση κοινών και μη EC numbers για N.gaditana και C.reinhardtii
(AlgaGEM). Τροποποιήθηκαν τα αρχεία έτσι ώστε οι κωδικοί των ενζύμων
να βρίσκονται στην 1η στήλη.
count=0;
count_miss=0;
for i=1:size(N_g,1) % σάρωσε την 1η στήλη του αρχείου excel του
N.gaditana, με όνομα N_g
check=0;
for j=1:size(C_r,1) % σάρωσε την 1η στήλη του αρχείου excel του
C.reinhardtii, με όνομα C_g
if findstr(char(C_r{j,1}),char(N_g{i,1}));
count=count+1;
% Αν βρεθούν κοινοί κωδικοί φτιάξε ένα νέο πίνακα κελιών (εδώ
ονομάστηκε common), όπου τα στοιχεία της πρώτης στήλης θα «γεμίζουν»
από την πρώτη στήλη του αρχείου του Nannochloropsis και οι υπόλοιπες
στήλες από το αρχείο του Chlamydomonas
common(count,2:size(C_r,2))=C_r(j,2:size(C_r,2));
common(count,1)=N_g(i,1);
check=1;
end
end
% Για τους κωδικούς των ενζύμων του Nannochloropsis που δεν βρέθηκαν
μέσα στο αρχείο του C.reinhardtii φτιάξε ένα νέο πίνακα κελιών (εδώ
ονομάστηκε missed), όπου όλες οι στήλες θα προέρχονται από το αρχείο
του Nannochloropsis.
if check==0
count_miss=count_miss+1;
missed(count_miss,:)=N_g(i,:);
end
end

```

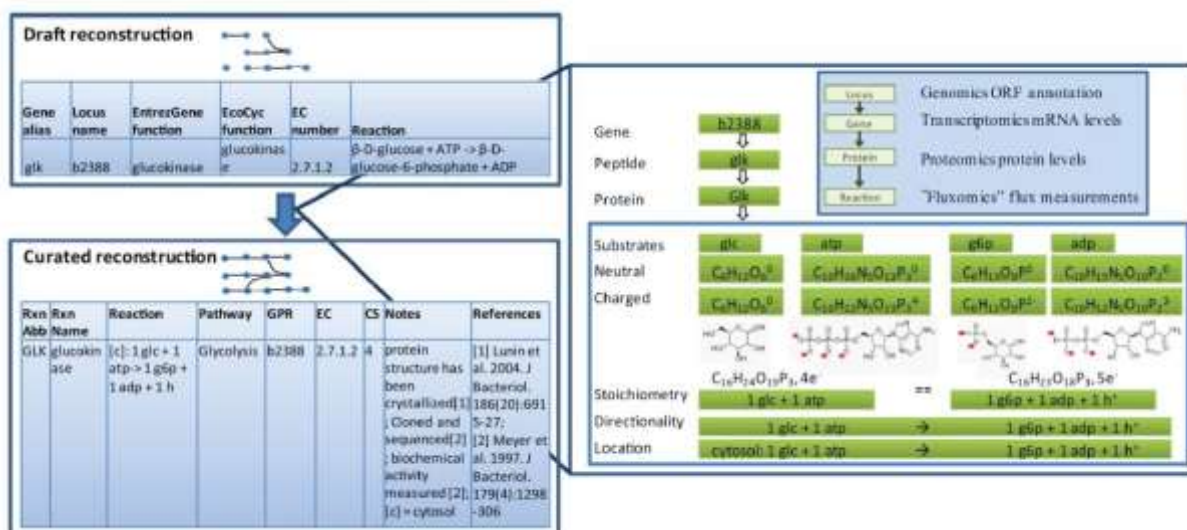
Για να λάβουμε τα αποτελέσματα της σύγκρισης σε αρχείο excel χρησιμοποιήσαμε την εντολή `xlswrite(filename,A)`. Τελικά βρέθηκαν 251 EC numbers που πρέπει να συμπληρωθούν.

2| Εύρεση υποψηφίων μεταβολικών αντιδράσεων για τα ένζυμα που βρέθηκαν.

Με συνδυαστική χρήση των βάσεων δεδομένων KEGG, Brenda και ExPASy-ENZYME πραγματοποιήθηκε αντιστοίχιση των EC numbers με τις αντιδράσεις που αυτά καταλύουν. Σημειώνεται ότι αυτό το στάδιο έχει ως στόχο την δημιουργία μιας λίστα υποψηφίων, τα οποία δεν θα είναι απαραίτητα ολοκληρωμένα ή περιεκτικά. Πολλά ψευδώς θετικά θα υπάρχουν στη λίστα. Για παράδειγμα, πρωτεΐνες που συμμετέχουν στη μεθυλίωση του DNA ή στην τροποποίηση του rRNA έχουν E.C. numbers, αλλά η δράση τους δεν λαμβάνεται υπόψιν.

3| Συναρμολόγηση του προσχεδίου.

Συλλογή όλων των υποψηφίων μεταβολικών γονιδίων και των αντιδράσεών τους σε ένα φύλλο excel. Αυτό το φύλλο θα είναι η αφετηρία της μη αυτοματοποιημένης (χειροκίνητης) βελτίωσης του προσχεδίου.



Εικόνα 5.1 Μη αυτοματοποιημένη (χειροκίνητη) βελτίωση του προσχεδίου. Το προσχέδιο μετατρέπεται σε βελτιωμένο προσχέδιο με επανεξέταση του περιεχομένου του. Συγκεκριμένα, οι μεταβολικές αντιδράσεις που έχουν αποκτηθεί από βιοχημικές βάσεις δεδομένων ή βιβλιογραφία ελέγχονται μία προς μία για την εισαγωγή τους στο μοντέλο, μέσω ισοζυγίων μάζας και φορτίου.

Τα πρώτα δύο στάδια κατασκευής του προσχεδίου πραγματοποιούνται σε φύλλα excel. Είναι σημαντικό η σειρά των στηλών στο φύλλο excel να έχει συγκεκριμένη μορφή καθώς και ονομασίες. Στις Εικ.2 και Εικ.3 δίνεται η μορφή των φύλλων excel για τις αντιδράσεις και τους μεταβολίτες.

A	B	C	D	E	F	G	H	I	J	K	L	M		
1	Required	Required	Required	Required	Optional	Required	Optional	Optional	Optional	Optional	Recommended	Recommended		
2	Rxn name	Rxn description	Formula	Gene-reaction association	Genes	Protein	Subsystem	Reversible	LB	UB	Objective	ConfidenceScore	E.C. number	Notes
		Fructose transport via PEP-Pyr PTS (f6p fru[e] + pep[c] -> f6p[c] + pyr[c])		(b1817 and b1818 and b1819 and b2415 and b2416)	b1817 b1818 b1819 b2415 b2416		Transport, Extracellular		0	0	1000	0	4 N/A	
3	FRUpts2		atp[c] + f6p[c] -> phosphofructokinase	(b3916 or b1723)	b1723 b3916		Glycolysis/ Gluconeogenesis		0	0	1000	0	4 2.7.1.11	E. coli has to gene pfkA
4	PFK		h[c]											PMID: 7004378 ; F phosphoglucose i respiratory contr
5	PGI	glucose-6-phosphate isomerase	g6p[c] <=> f6p[c]	b4025 ((b2417 and b1101 and b2415 and b2416) or (b1817 and b1818 and b1819 and b2415 and b2416) or (b2417 and b2416 b2417	b4025 b1101 b1621 b1817 b1818 b1819 b2415 b2416 b2417		Glycolysis/ Gluconeogenesis	1	-1000	1000	0	4 5.3.1.9	the enzyme (both parallelly with oth	
6	GLCpts	D-glucose transport via PEP-Pyr PTS	glc-D[e] + pep[c] -> g6p[c] + pyr[c]				Transport, Extracellular		0	0	1000	0	4	

Εικόνα 5.2 Υπόδειγμα φύλλου excel για τη συλλογή αντιδράσεων.

A	B	C	D	E	F	G	H	I	J	K	L
1	Required	Required	Required	Required	Optional	Recommended	Recommended	Recommended	Recommended	Recommended	Recommended
2	Metabolite name	Metabolite description	Neutral metabolite formula	Charged metabolite formula	Metabolite charge	Metabolite Compartment	Metabolite KeggID	Metabolite PubChemID	Metabolite Inchi string	Metabolite Smiles	
3	adp[c]	ADP	C10H15N5O10P2	C10H12N5O10P2	-3	cytosol	C00008	3310	16761-1/C10H15 N4 ncnc2n cnc12 C@		
4	atp[c]	ATP	C10H16N5O13P3	C10H12N5O13P3	-4	cytosol	C00002	3304	15422-1/C10H16 N4 ncnc2n cnc12 C@		
5	fbp[c]	D-Fructose-6-p	C6H13O9P	C6H11O9P	-2	cytosol	C00065	3385	15946-1/C6H13O [H][C@](O)(COP)		
6	fdp[c]	D-Fructose-1-6	C6H14O12P2	C6H10O12P2	-4	cytosol	C00354	3647	16905-1/C6H14O [H][C@](O)(COP)(COP)		
7	fru[e]	D-Fructose	C6H12O6	C6H12O6	0	cytosol	C00095	3395	15824-1/C6H12O C(C1C(C(C1O))CO)O		
8	g6p[c]	D-Glucose-6-p	C6H13O9P	C6H11O9P	-2	cytosol	C00092	3392	4170-1/C6H13O OC1O[C@H](COP)(COP)		
9	glc-D[e]	D-Glucose	C6H12O6	C6H12O6	0	extracellular	C00031	3333	4167-1/C6H12O OC(C@H)OC(O)C@H		
10	h[c]	H	H	H	1	cytosol	C00080	3380	15378-1/p+1/fh/ [H+]		
11	pep[c]	Phosphoenolp	C3H5O6P	C3H2O6P	-3	cytosol	C00074	3374	18021-44897-1/C3H5O6OP(O)=O)OC=C(C)O		
12	pyr[c]	Pyruvate	C3H4O3	C3H3O3	-1	cytosol	C00022	3324	32816-1/C3H4O3 CC(=O)C(O)=O		
13											

Εικόνα 5.3 Υπόδειγμα φύλλου excel για τη συλλογή μεταβολιτών.

4| Συλλογή πειραματικών δεδομένων.

Η διαδικασία της χειροκίνητης επιμέλειας εξαρτάται σε μεγάλο βαθμό από πειραματικά, ειδικά για τον οργανισμό δεδομένα. Συγκεντρώνονται όλες οι πιθανές πληροφορίες.

2^ο ΣΤΑΔΙΟ : ΜΗ ΑΥΤΟΜΑΤΟΠΟΙΗΜΕΝΗ (ΧΕΙΡΟΚΙΝΗΤΗ) ΒΕΛΤΙΩΣΗ-ΕΠΙΜΕΛΕΙΑ ΤΟΥ ΠΡΟΧΕΔΙΟΥ (Manual reconstruction refinement)

5| Προσδιορισμός και επαλήθευση χρήσης υποστρώματος και συμπαράγοντα.

Το προσχέδιο του μοντέλου περιλαμβάνει ένα σύνολο μεταβολικών γονιδίων και καταλυτικών δράσεων που πιστεύεται ότι υπάρχουν στον οργανισμό του *Nannochloropsis*. Λόγω του ότι μπορεί να υπάρχουν πιθανά λάθη, κάθε καταχώριση θα πρέπει να υποστηριχθεί από πειραματικά δεδομένα ή βιβλιογραφία. Όσες από τις 251 αντιδράσεις έχουν συλλεχθεί, αφού ελεγχθούν, θα συνενωθούν με το μοντέλο του *Chlamydomonas reinhardtii*, το οποίο είναι φυλλογενετικά συγγενικό είδος. Θα πρέπει να ελεγχθούν και ποιες αντιδράσεις που περιέρονται σε αυτό το μοντέλο είναι παρούσες στον *Nannochloropsis*.

Επιπλέον, σε κάποιες περιπτώσεις ενδείκνυται να αποκλείονται συγκεκριμένες αντιδράσεις από το να συμπεριληφθούν στο μοντέλο. Τέτοιες αντιδράσεις, οι οποίες περιέχουν γενικούς όρους, όπως πρωτεΐνη (protein), DNA, αποδέκτης ηλεκτρονίων (electron acceptor) δεν θα πρέπει να εισαχθούν καθώς δεν είναι αρκετά ειδικές και συνήθως βρίσκονται στις βάσεις δεδομένων έως ότου γίνουν διαθέσιμα περισσότερα βιοχημικά δεδομένα. Βάσει αυτής της λογικής, από τις 251 αντιδράσεις που ανακτήσαμε παραμείναν οι 229.

Μέσω στοιχειώδους βιβλιογραφίας, και με λιγότερη χρήση των βάσεων δεδομένων (BΔ) KEGG και Brenda, ορίζεται και επιβεβαιώνεται η εξειδίκευση του υποστρώματος και του συμπαράγοντα του κάθε ενζύμου για τον οργανισμό-στόχο. Ένας γενικός κανόνας είναι κανείς να υποθέσει ότι τα ένζυμα, τα οποία σχετίζονται με μία αντίδραση στην BΔ KEGG για

παράδειγμα, δεν χρειάζονται ειδικό έλεγχο για το αν είναι εξειδικευμένα για τον οργανισμό.

ΠΡΟΣΟΧΗ Συχνά μόνο βιοχημικά δεδομένα μπορούν να αποκαλύψουν το σωστό συμπαράγοντα και υπόστρωμα καθώς η θέση προσκόλλησης μπορεί να μην είναι διακριτή στη γονιδιακή αλληλουχία συναφών μεταβολιτών.

Ένα βασικό κομμάτι της κατασκευής του μοντέλου είναι ο τρόπος γραφής των αντιδράσεων και των συντομογραφιών. Στο μοντέλο του *Nannochloropsis* ακολουθήθηκαν συγκεκριμένοι κανόνες σύνταξης των αντιδράσεων και οι οποίοι υποδεικνύονται στον πίνακα Π.Ι.1 του Παραρτήματος. Οι μεταβολίτες που έχουμε συλλέξει και θα συλλέγουμε στο επόμενο στάδιο προέρχονται από τη βάση KEGG και έχουν συγκεκριμένη ονομασία. Για να εισαχθούν στο μοντέλο υπάρχει συγκεκριμένος τρόπος γραφής τους βασιζόμενοι στην ονομασία τους από τη βάση δεδομένων. Οι κανόνες των συντομογραφιών παρουσιάζονται στον πίνακα Π.Ι.2 του Παραρτήματος.

6| Εύρεση ουδέτερου τύπου για κάθε μεταβολίτη σε κάθε αντίδραση.

Χρησιμοποιήθηκε η ΒΔ KEGG, ταυτόχρονα με τη διαδικασία εύρεσης αντιδράσεων που αντιστοιχούν σε κάθε EC number. Παρόλο που η βάση αυτή δεν είναι τόσο περιεκτική είναι η πιο προσιτή. Έγινε επαλήθευση μέσω της ΒΔ Brenda.

7| Υπολογισμός στοιχειομετρίας αντίδρασης.

Αν και οι αντιδράσεις που λήφθηκαν από την KEGG είχαν ήδη στοιχειομετρικούς συντελεστές έγινε επαλήθευση με μέτρηση κάθε στοιχείου σε κάθε πλευρά της εξίσωσης. Το γεγονός ότι δεν υπήρχαν φορτισμένοι μεταβολίτες διευκόλυνε τη διαδικασία.

8| Προσδιορισμός κατεύθυνσης αντίδρασης

Έγινε με συνδυαστική χρήση βιοχημικών δεδομένων, βιβλιογραφίας και μεταβολικών χαρτών της KEGG και MetaCyc. Αν δεν υπήρχαν δεδομένα ακολουθήθηκε ο εξής γενικός κανόνας: 1) αντιδράσεις που αφορούν μεταφορά φωσφορικού άλατος από ATP σε ένα μόριο αποδέκτη θα είναι

αναντίστρεπτες. (με εξαίρεση της ATP συνθάσης, η οποία είναι αντιστρεπτή).

2) αντιδράσεις που περιέχουν κινόνες είναι γενικά αναντίστρεπτες.

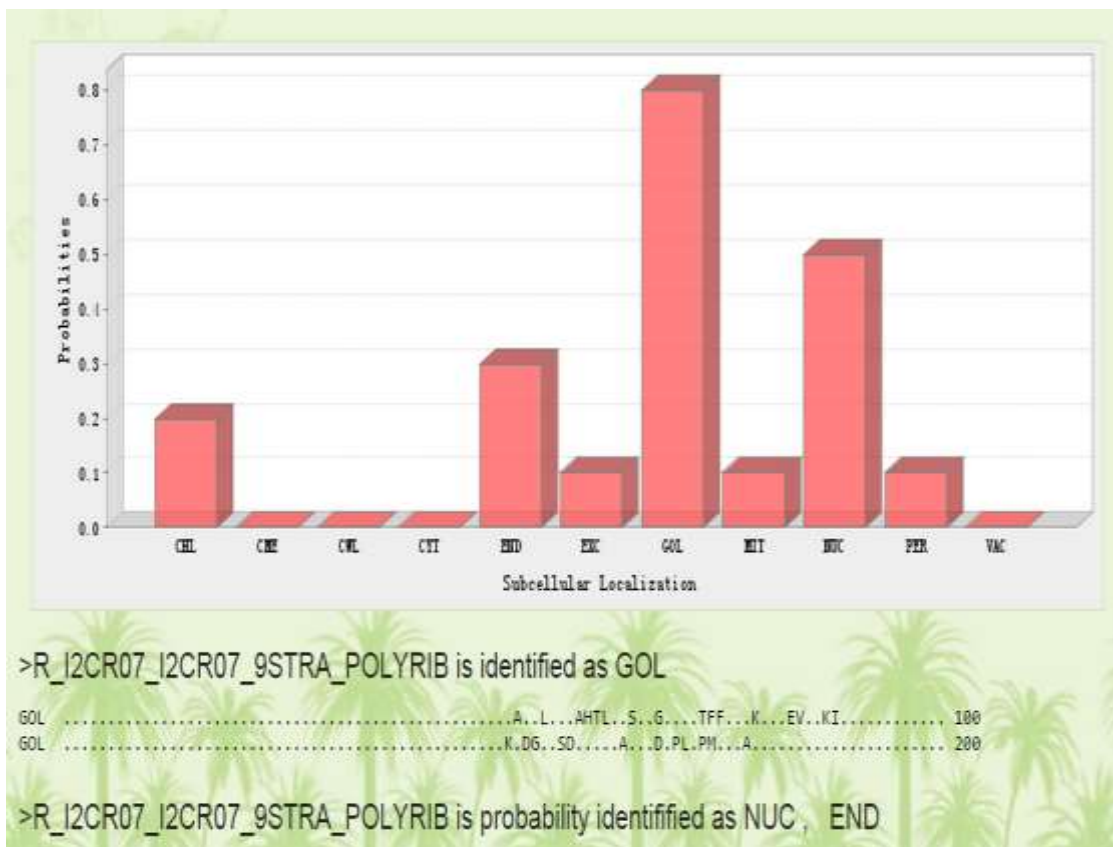
Σημειώνεται ότι είναι πολύ σημαντική η ανάθεση σωστής κατεύθυνσης καθώς λάθος κατεύθυνση σε μια αντίδραση έχει σημαντική επίδραση στην επίδοση του μοντέλου.

9| Προσθήκη πληροφοριών εντοπισμού (localization) για την αντίδραση και το γονίδιο.

Το μοντέλο που κατασκευάστηκε έχει τα εξής κυτταρικά διαμερίσματα: κυτταρόπλασμα (cytoplasm), μιτοχόνδριο (mitochondrion), χλωροπλάστης (chloroplast), υπεροξειδισμόμα (peroxisome), και περίπλασμα (periplasm). Χρησιμοποιήθηκε η BD Uniprot καθώς και ειδικοί οι αλγόριθμοι WolfPSORT, PlantLoc, iLoc-Euk, Euk-mPloc και Plant-mPloc όταν δεν υπήρχαν πειραματικά δεδομένα. Οι αλγόριθμοι αυτοί βρίσκονται διαδικτυακά και είναι ελεύθερα προσβάσιμοι. Χρησιμοποιούνται για την πρόβλεψη του κυτταρικού διαμερίσματος πρωτεϊνών βασιζόμενοι στην αλληλουχία βάσεων του κάθε γονιδίου. Στις Εικ.4 και Εικ.5 φαίνεται η χρήση ενός από αυτούς τους αλγορίθμους και η μορφή των αποτελεσμάτων.



Εικόνα 5.4 Χρήση αλγορίθμου PlantLoc για τον εντοπισμό του κυτταρικού διαμερίσματος ενός γονιδίου



Εικόνα 5.5 Αποτελέσματα του αλγορίθμου σε πιθανότητες. Στο συγκεκριμένο παράδειγμα βλέπουμε ότι η αντίδραση που αντιστοιχεί σε αυτό το γονίδιο έχει μεγαλύτερες πιθανότητες να συμβαίνει στο σύστημα Golgi και μετά στον πυρήνα.

Για κάθε αντίδραση που βρέθηκε το κυτταρικό διαμέρισμα ελέγχθηκε η σύνδεση της με το ήδη υπάρχων μεταβολικό μοντέλο για να μη δημιουργηθούν περαιτέρω κενά.

10| Προσθήκη υποσυστήματος (μονοπατιού) για την αντίδραση.

Αυτή η διαδικασία είναι πολύ σημαντική και μπορεί να βοηθήσει στο στάδιο του εντοπισμού σφαλμάτων και αξιολόγησης του δικτύου. Βασιστήκαμε στην ΒΔ KEGG και σε μεταβολικούς χάρτες για τη συμπλήρωση των μονοπατιών στα οποία συμμετέχει η αντίδραση.

11| Επαλήθευση σχέσης γονιδίου-πρωτεΐνης-αντίδρασης (GPR).

Προσδιορίζεται αν η λειτουργική πρωτεΐνη είναι ετερομερικό ενζυμικό σύμπλοκο. Δηλαδή αν το ένζυμο μπορεί να καταλύσει παραπάνω από μία αντίδραση και αν παραπάνω από ένα ένζυμο μπορεί να καταλύσει την ίδια αντίδραση. Χρησιμοποιήθηκε η ΒΔ KEGG για τη συμπλήρωση αυτής της πληροφορίας. Η σχέση γονιδίου-πρωτεΐνης-αντίδρασης αντανακλά την απαίτηση για όλα τα γονίδια μέσα στην αντίδραση με χρήση του κανόνα Boolean AND, αλλιώς με χρήση του κανόνα OR. Για παράδειγμα:

(gene1 and gene2): απαιτούνται και τα δύο αυτά γονίδια ή

(gene1 or gene 2): απαιτείται ένα από τα δύο γονίδια

12| Επαναλαμβάνονται τα βήματα 5 έως 11 για όλα τα ένζυμα (συνεπώς γονίδια και αντιδράσεις) του προσχεδίου.

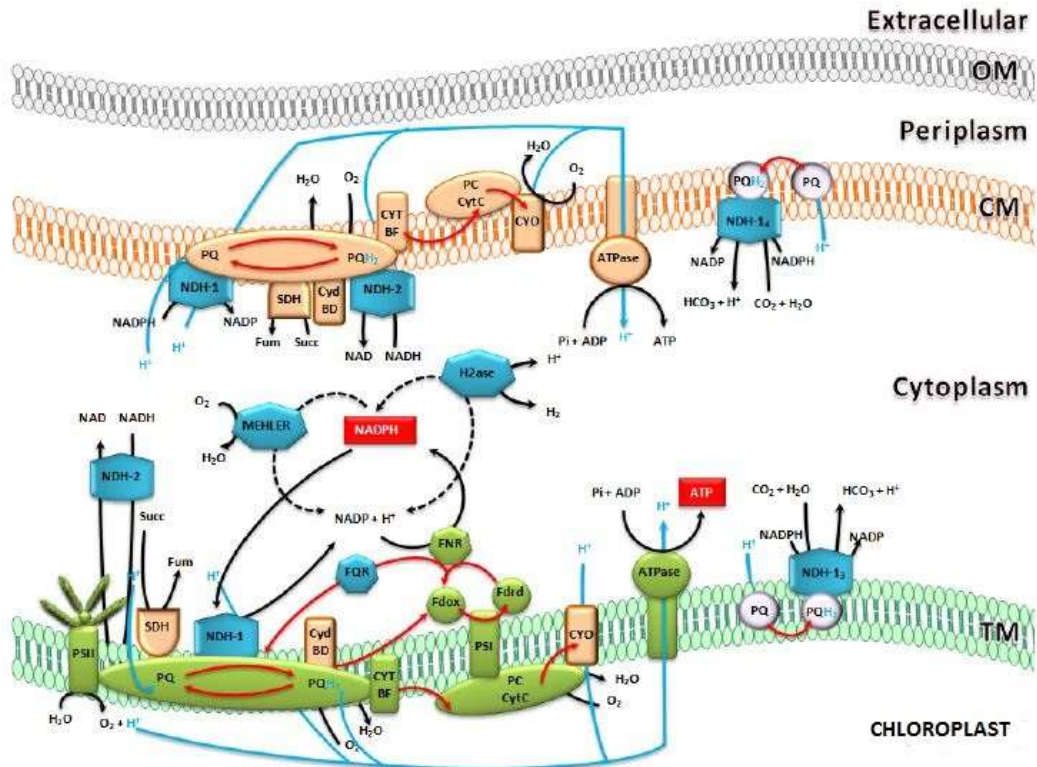
(Thiele & Palsson, 2010)

13| *Κατασκευή αντιδράσεων φωτοσύνθεσης και οξειδωτικής φωσφορυλίωσης.

Το μικροφύκος *Nannochloropsis gaditana* είναι ένας φωτοσυνθετικός οργανισμός. Το μοντέλο του AlgaGEM δεν καλύπτει το μονοπάτι της φωτοσύνθεσης (καλύπτει μερικώς τη φωτοαναπνοή) ενώ παράλληλα δίνει δύο συνολικές αντιδράσεις για να περιγράψει το μονοπάτι της οξειδωτικής φωσφορυλίωσης. Για τη μοντελοποίηση τους χρησιμοποιήθηκαν αντιδράσεις από το μοντέλο του *Synechocystis*, iJN678, το οποίο είναι ένα φωτοσυνθετικό

βακτήριο. Μοντελοποιήθηκαν λεπτομερώς και η οξειδωτική φωσφορυλίωση και ο μηχανισμός της φωτοσύνθεσης.

Θεωρήθηκε –προς διευκόλυνση της σύνδεσης με το υπόλοιπο μοντέλο- ότι οι φωτοσυνθετικές και αναπνευστικές αλυσίδες μεταφοράς ηλεκτρονίων περιέχονται στον χλωροπλάστη και όχι στη θυλακοειδή μεμβράνη όπως στο *Synechosystis*, ενώ μόνο η αναπνευστική αλυσίδα είναι παρούσα στο κυτταρόπλασμα. Το μονοπάτι της οξειδωτικής φωσφορυλίωσης περιλαμβάνει τα εξής σύμπλοκα NAD(P)H αφυδρογονασών: NDH-1, NDH-2, NDH-1₃ και NDH-1₄, το σύμπλοκο αφυδρογονάσης του ηλεκτρικού άλατος SDH, δύο aa3-τύπου τερματικές οξειδάσες (CtaI και CtaII, CYO) και την οξειδάση της πλαστοκινόνης (CydBD) μαζί με σύμπλοκα αδενοσινοτριφωσφατάσης (ATPase). Η φωτοσυνθετική αλυσίδα ηλεκτρονίων περιλαμβάνει το φωτοσύστημα II (PSII), το κυτόχρωμα c, το φωτοσύστημα I (PSI) και την NADP⁺ οξειδοαναγωγή ως συνιστώσες της γραμμικής αλυσίδας ηλεκτρονίων (linear electron chain). Επιπλέον συμπεριλήφθηκε και μία αντίδραση αναγωγής της πλαστοκινόνης καθώς και βοηθητικά μονοπάτια φωτοσυνθετικής μεταφοράς ηλεκτρονίων, όπως η αντίδραση Mehler. (Montagud, et al., 2010)

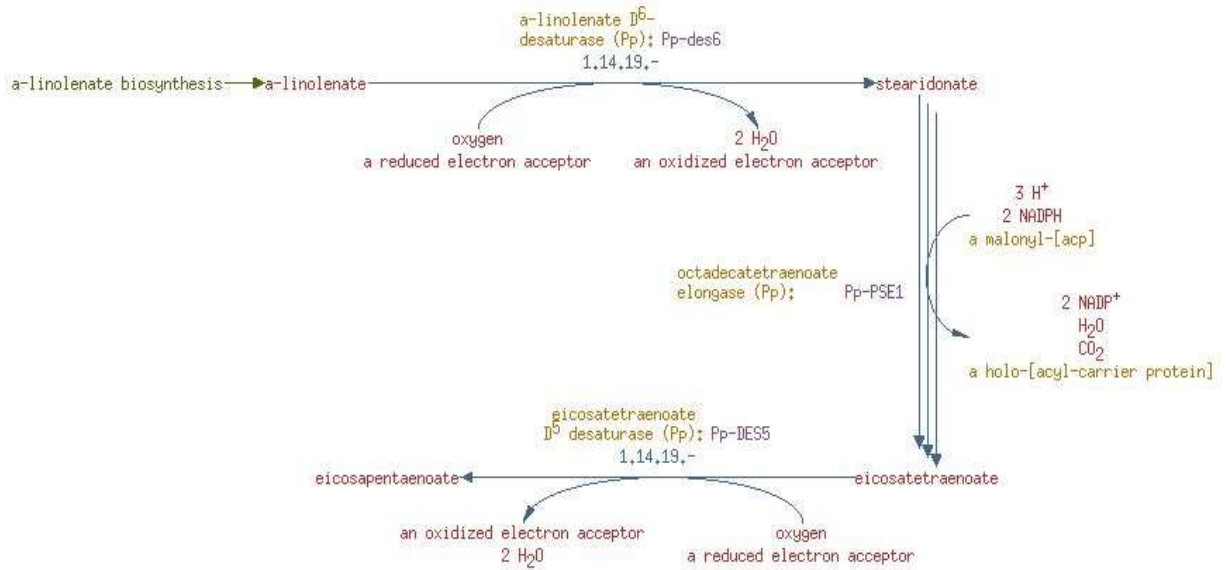


Εικόνα 5.6 Μοντελοποίηση της οξειδωτικής φωσφορυλίωσης και φωτοσύνθεσης στο μοντέλο του *Nannochloropsis*. Στην κατασκευή περιλαμβάνονται τα κυτταρικά διαμερίσματα: εξωκυτταρικό, περίπλασμα, κυτταρόπλασμα και χλωροπλάστης, παράλληλα στην εικόνα φαίνονται η εξωτερική, κυτταροπλασματική και θυλακοειδή μεμβράνη, οι οποίες προς διευκόλυνση δεν έχουν εισαχθεί στο μοντέλο. Η φωτοσυνθετική ροή ηλεκτρονίων (πράσινο χρώμα) και το μονοπάτι της αναπνοής (πορτοκαλί) καθώς και οι βοηθητικές διακλαδώσεις αυτών (με μπλε χρώμα). Συντομογραφίες: SDH: σύμπλοκο αφυδρογονάσης του ηλεκτρικού οξέος; PQ: πλαστοκινόνη; CYTBF: κυτόχρωμα Cytb₆f; PC: πλαστοκυανίνη; CytC: cytochrome c; CYO: οξειδάση του κυτοχρώματος; CYTBD: οξειδάση πλαστοκινόνης; Fdox: οξειδωμένη ferredoxin; Fdrd: ανηγμένη ferredoxin; FNR: ferredoxin NADP+ αναγωγίαση; FQR: ferredoxin αναγωγίαση πλαστοκινόνης; H₂ase: αμφίδρομη υδρογονάση. Η ροή των ηλεκτρονίων και πρωτονίων μέσω των φωτοσυνθετικών και αναπνευστικών μονοπατιών υποδεικνύεται από τα κόκκινα και μπλε βέλη αντίστοιχα. Τα τελικά προϊόντα της φωτοσύνθεσης, δηλαδή ATP και NADPH υποδεικνύονται με κόκκινα κουτάκια.

14| *Κατασκευή αντιδράσεων σύνθεσης κορεσμένων και πολυακόρεστων λιπαρών οξέων.

Η σύνθεση λιπαρών οξέων στο *Nannochloropsis* κυριαρχείται από το παλμιτικό οξύ C16:0, το παλμιτολεϊκό οξύ C16:1 και το εικοσαπεντανοϊκό οξύ C20:5(n-3), τα οποία μαζί αποτελούν περισσότερο του 60% της ολικής μάζας λιπαρών οξέων. Αυτό καθιστά σημαντική την μοντελοποίηση της σύνθεσης αυτών των οξέων και κυρίως των πολυακόρεστων. Για την κατασκευή αυτών των αντιδράσεων χρησιμοποιήθηκε η ΒΔ Kegg – η οποία όμως είναι ελλιπής, ειδικά για τη σύνθεση του εικοσαπεντανοϊκού- και κυρίως η MetaCyc, μία βάση δεδομένων που περιλαμβάνει μεταβολικούς χάρτες σχεδιασμένους βάσει πειραματικών δεδομένων. Το μοντέλο του AlgaGEM καλύπτει τη σύνθεση του παλμιτικού οξέος. Με τη βοήθεια των μεταβολικών χαρτών της Kegg συμπληρώθηκαν οι αντιδράσεις σύνθεσης των κορεσμένων λιπαρών οξέων μυριστικού και στεατικού, ενώ με συνδυαστική χρήση Kegg-MetaCyc οι αντιδράσεις σύνθεσης του παλμιτολεϊκού, ολεϊκού, λινολεϊκού, α-λινολενικού, γ-λινολενικού, αραχιδονικού και εικοσαπεντανοϊκού. Η βιοσύνθεση του αραχιδονικού ξεκίνησε από το γ-λινολενικό (συγκεκριμένα από το g-linolenoyl-CoA), ενώ του εικοσαπεντανοϊκού από το α-λινολενικό (α-linolenoyl-CoA) και περιλαμβάνει διαδοχικά Δ6 αποκορεσμό, επιμήκυνση και Δ5 αποκορεσμό. (Sukenik, 1999) (Caspi & SRI International, 2011)

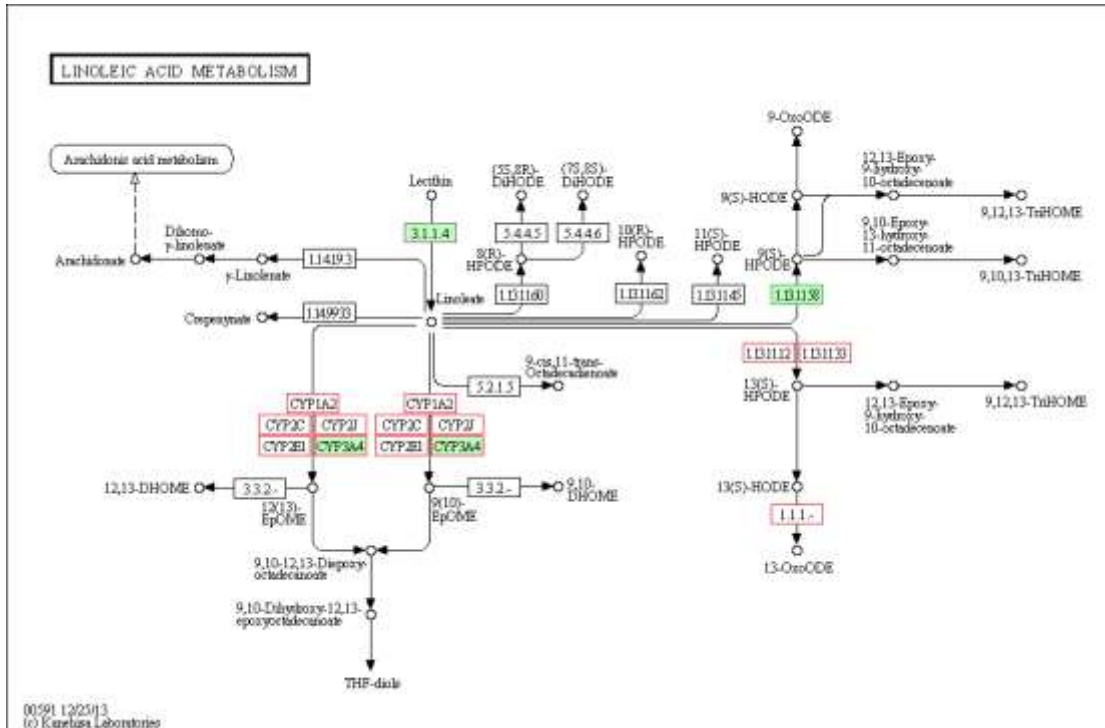
Παρακάτω παρατίθεται το μονοπάτι της σύνθεσης του εικοσαπεντανοϊκού οξέος.



Εικόνα 5.7 Μονοπάτι βιοσύνθεσης εικοσαπεντανοϊκού οξέος για μικροφύκη και κατώτερα φυτά, από τη βάση MetaCyc.

15| Κατασκευή μεταβολικών χαρτών.

Μέσω της ΒΔ KEGG προμηθευτήκαμε τους μεταβολικούς χάρτες του *Nannochloropsis* και οι οποίοι συμπληρώθηκαν βάσει του ήδη υπάρχοντος μοντέλου αλλά και των αντιδράσεων που προστέθηκαν. Στο Παράρτημα II βρίσκονται οι πίνακες του βασικού μεταβολισμού που καλύπτει το μοντέλο του *Nannochloropsis*. Ένα παράδειγμα τέτοιου χάρτη δίνεται στην Εικ.8.



Εικόνα 5.8 Μονοπάτι του μεταβολισμού λινολεϊκού οξέος. Τα κουτιά με πράσινο χρώμα αντιπροσωπεύουν ένζυμα που είναι εξειδικευμένα για το *Nannochloropsis*. Τα κουτιά που έχουν κόκκινο περίγραμμα προέρχονται από το υπάρχον μοντέλο AlgGEM, με πράσινο περίγραμμα τα ένζυμα για τα οποία έχει βρεθεί γονίδιο του *Nannochloropsis* και με μωβ περίγραμμα τα ένζυμα για τα οποία δεν βρέθηκε γονίδιο του *Nannochloropsis* (orphan reactions) και συμπληρώθηκαν κατά το στάδιο συμπλήρωσης κενών (gap filling).

16| Προσθήκη αυθόρμητων αντιδράσεων.

Προστέθηκαν μόνο εκείνες οι αντιδράσεις οι οποίες έχουν τουλάχιστον ένα μεταβολίτη παρών στο μοντέλο ή για τη συμπλήρωση κενών στο δίκτυο, προκειμένου να μειωθεί ο αριθμός των αδιέξοδων μεταβολιτών.

17| Προσθήκη εξωκυτταρικών και περιπλασμικών αντιδράσεων μεταφοράς στο μοντέλο.

Αυτή η διαδικασία έγινε κυρίως κατα το στάδιο εντοπισμού σφαλμάτων για να ενισχύσει τη συνδετικότητα του δικτύου. Ο κανόνας είναι ότι για κάθε μεταβολίτη που είναι γνωστό ότι λαμβάνεται από το μέσο καλλιέργειας ή είναι γνωστό ότι εκκρίνεται στο μέσο, θα πρέπει να υπάρχει μια αντίδραση μεταφοράς (από τον εξωκυτταρικό χώρο στο κυτταρόπλασμα/περίπλασμα).

18| Προσθήκη αντιδράσεων ανταλλαγής (exchange reactions).

Τέτοιες αντιδράσεις προστίθενται για όλους τους εξωκυτταρικούς μεταβολίτες και αντιπροσωπεύουν τα όρια του συστήματος.

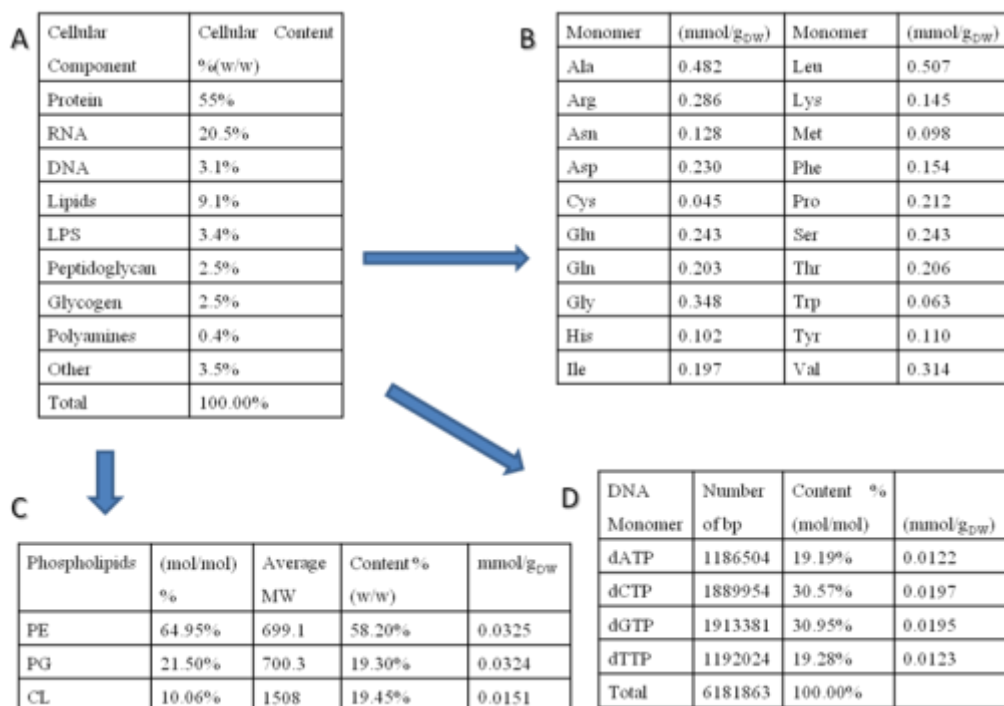
19| Προσθήκη ενδοκυτταρικών αντιδράσεων μεταφοράς.

Προστέθηκαν μόνο εκείνες οι αντιδράσεις που χρειαζόντουσαν για το κλείσιμο «κενών».

ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΣΥΣΤΑΣΗΣ ΒΙΟΜΑΖΑΣ

20| Προσδιορισμός της χημικής σύστασης του κυττάρου του *Nannochloropsis* σε πρωτεΐνες, RNA, DNA, λιπίδια κ.α..

Για τις πληροφορίες αυτές χρησιμοποιήθηκε εκτεταμένη βιβλιογραφία λόγω έλλειψης πειραματικών δεδομένων.



Εικόνα 5.9 Προσδιορισμός χημικής σύστασης του κυττάρου. Η συνεισφοράς των μονομερών στοιχείων είναι σε mmol/g_{DW}.

Στον Πίνακα 5.1 παρατίθεται η μακρομοριακή κυτταρική σύσταση του *Nannochloropsis* σε γραμμάρια συνιστώσας ανά γραμμάρια ξηρής μάζας κυττάρων. (Brown, 1991)

Πίνακας 5.1 Μακρομοριακή κυτταρική σύσταση του *Nannochloropsis* (Brown, 1991)

Ποσοστό Ξηρής μάζας κυττάρων(ξ.μ.κ. ή DW) [% g /g ξ.μ.κ.]	
Τέφρα	8,66
Χλωροφύλλη α	6
Υγρασία	1,95
Πρωτεΐνη	35
Υδατάνθρακες	7,8
Λιπίδια	18
DNA	0,61
RNA	1,71
Καροτενοειδή	6,6
Ολικό ποσοστό	86,33

21| Προσδιορισμός σύστασης σε αμινοξέα.

Για τον προσδιορισμό της σύστασης σε αμινοξέα (πρωτεΐνες) βρέθηκαν δεδομένα βιβλιογραφικά. Το πρωτεϊνικό κλάσμα αντιπροσωπεύει όλες τις κυτταρικές πρωτεΐνες. Η σύσταση των αμινοξέων του *Nannochloropsis gaditana* προσεγγίστηκε από δεδομένα για το *Nannochloropsis oculata*, συγγενικό είδος του *N. gaditana* και φυλογενετικά όμοιο σε ποσοστό μεγαλύτερο του 95%. (Brown, 1991)

Πίνακας 5.2 Σύσταση σε αμινοξέα του *Nannochloropsis* (Brown, 1991)

% κ.β. [g αμινοξέος/g πρωτεΐνη]

Ισολευκίνη	4,8
Βαλίνη	6,5
Αλανίνη	7,4
Προλίνη	9,3
Τυροσίνη	4,2
Ιστιδίνη	2,1
Φαινυλαλανίνη	6,2
Ασπαραγινικό οξύ	7,6
Θρεονίνη	5,5
Σερίνη	5
Γλουταμινικό οξύ	10,1
Γλυκίνη	5,5
Κυστεΐνη	0,41
Μεθειονίνη	1,6
Λευκίνη	7,8
Λυσίνη	6,1
Αργινίνη	7,3
Τρυπτοφάνη	1,6
γ-αμινοβουτυρικό οξύ	0,78
Υδροξυπρολίνη*(είναι μη πρωτεϊνογενετικό αμινοξύ, δεν συμπεριλήφθηκε στο μοντέλο)	0,05
Ορνιθίνη	0,35

Με χρήση του ποσοστού κατά βάρος καθώς και του μοριακού βάρους κατάλοιπου αμινοξέος (αφού ο πολυμερισμός συνήθως σχετίζεται με την απώλεια ενός μορίου νερού) υπολογίζουμε τα mole αμινοξέος ανά gr ολικής πρωτεΐνης. Στη συνέχεια πολλαπλασιάζουμε με τα γραμμάρια ολικής πρωτεΐνης ανά γραμμάρια ξηρής μάζας κυττάρου (έχουμε το ποσοστό κατά βάρος από τη μακρομοριακή σύσταση) και βρίσκουμε τα mole αμινοξέος ανά gr ξηρής μάζας κυττάρου. Με μετατροπές μονάδων καταλήγουμε σε mmol αμινοξέος/g ξ.μ.κ, που θα αποτελέσει το

στοιχειομετρικό συντελεστή του κάθε αμινοξέος στην συνολική αντίδραση-εξίσωση της βιομάζας. Στον Πίνακα 5.3 δίνονται οι μετατροπές και οι τελικοί στοιχειομετρικοί συντελεστές.

Πίνακας 5.3 Υπολογισμοί κλασματικής κατανομής αμινοξέων στη βιομάζα σε mmol/g ξηρής μάζας κυττάρων

	MB κ.αμινοξέος	[mol αμινοξέος/g ολικής πρωτεΐνης]	[mmol αμινοξέος/g ξ.μ.κ.]
Ισολευκίνη	113,16	0,000424178	0,148
Βαλίνη	99,13	0,000655705	0,229
Αλανίνη	71,08	0,00104108	0,364
Προλίνη	97,12	0,000957578	0,335
Τυροσίνη	163,18	0,000257384	0,090
Ιστιδίνη	137,14	0,000153128	0,054
Φαινυλαλανίνη	147,18	0,000421253	0,147
Ασπαραγινικό οξύ	115,09	0,000660353	0,231
Θρεονίνη	101,1	0,000544016	0,190
Σερίνη	87,08	0,000574185	0,201
Γλουταμινικό οξύ	129,11	0,000782279	0,274
Γλυκίνη	57,05	0,000964067	0,337
Κυστεΐνη	103,14	3,97518E-05	0,014
Μεθειονίνη	131,2	0,000121951	0,043
Λευκίνη	113,16	0,00068929	0,241
Λυσίνη	128,17	0,00047593	0,167
Αργινίνη	156,19	0,000467379	0,164
Τρυπτοφάνη	186,21	8,59245E-05	0,030
γ-αμινοβουτυρικό οξύ	85,12	9,16353E-05	0,032
Υδροξυπρολίνη	113,13	4,41969E-06	0,002
Ορνιθίνη	114,16	3,06587E-05	0,011

22| Προσδιορισμός σύστασης νουκλεοτιδίων και ριβονουκλεοτιδίων.

Η σύσταση για το DNA και το RNA παρουσιάζεται στους Πίνακες 5.4 και 5.5.

Για το DNA τα δεδομένα λήφθηκαν από τη βάση δεδομένων, που περιέχει δεδομένα χλωροπλαστικών και μιτοχονδριακών γονιδιωμάτων για διάφορους οργανισμούς. Τα δεδομένα αυτά αφορούν το περιεχόμενο σε DNA του χλωροπλάστη και του μιτοχονδρίου του *Nannochloropsis* οπότε λόγω έλλειψης δεδομένων έγινε η υπόθεση ότι αντιπροσωπεύουν το περιεχόμενο σε DNA ενός κυττάρου του φύκου. (Sablok, et al., 2013)

Πίνακας 5.4 Σύσταση νουκλεοτιδίων-DNA (Sablok, et al., 2013)

	%(mol/mol DNA)	MB κατάλοιπου DNA [g/mole]
dATP	32,91	331,2
dTTP	34,14	304,2
dCTP	16,47	289,2
dGTP	16,48	304,2
Μέσο MB DNA [g/mole]		310,6152

Υπολογίζεται το μέσο μοριακό βάρος του DNA σε [g/mole]. Στη συνέχεια διαιρείται το μοριακό κλάσμα κάθε συστατικού με 100 και με το μέσο MB του DNA και έτσι βρίσκουμε τα mole ανά γραμμάριο DNA. Πολλαπλασιάζοντας με τα γραμμάρια DNA ανά γραμμάρια ξηρής μάζας κυττάρου και με κατάλληλες μετατροπές μονάδων καταλήγουμε σε mmol/g ξ.μ.κ. .

Πίνακας 5.5 Υπολογισμοί κατανομής νουκλεοτιδίων στη βιομάζα σε mmol/g ξ.μ.κ.

	mmol νουκλεοτιδίου/g ξ.μ.κ
dAMP	0,006
dTMP	0,007
dCMP	0,003
dGMP	0,003

Για το RNA δυστυχώς δεν βρέθηκαν δεδομένα ούτε βιβλιογραφικά ούτε σε βάσεις δεδομένων, οπότε ακολουθήθηκε διαφορετική πορεία για υπολογισμό του κλάσματος RNA στη βιομάζα. Βάσει βιβλιογραφίας η σύσταση ενός τυπικού ευκαρυωτικού κυττάρου είναι ~80% rRNA, ~15% tRNA και το υπόλοιπο είναι mRNA. (Lodish, et al., 2000) Η ριβονουκλεοτιδική σύσταση του mRNA θεωρήθηκε ίδια με του DNA. Η ριβονουκλεοτιδική σύσταση του rRNA υπολογίστηκε από τις αλληλουχίες των ριβοσωμικών μονάδων RNA 16S,18S. Η σύσταση του tRNA βρέθηκε από αλληλουχίες μεταφορέων RNA της προλίνης, αργινίνης και της θρεονίνης. Όλες οι αλληλουχίες λήφθηκαν από τη GenBank (National Center for Biotechnology Information, 2014)

Για το rRNA, έχουμε τον αριθμό της κάθε βάσης στο τμήμα 18S και 16S και υπολογίζουμε το κλάσμα της στο σύνολο του τμήματος, όπως φαίνεται στον Πίνακα 5.6. Ο μέσος όρος των δύο αυτών ριβοσωμικών τμημάτων μας δίνει τη σύσταση του rRNA.

Πίνακας 5.6 Υπολογισμοί μολαρικών κλασμάτων για το rRNA

Αζωτούχες βάσεις	Αριθμός βάσεων στο 18S	Μολαρικό κλάσμα βάσης	Αριθμός βάσεων στο 16S	Μολαρικό κλάσμα βάσης
A	475	0,265	477	0,265
G	476	0,266	479	0,267
U	480	0,268	481	0,268
C	359	0,201	360	0,200
	1790		1797	

Για το tRNA, έχουμε τον αριθμό της κάθε βάσης στον RNA μεταφορέα της προλίνης, της αργινίνης και της θρεονίνης (Πίνακας 5.7) και υπολογίζουμε το κλάσμα της στο σύνολο του τμήματος (Πίνακας 5.8). Ο μέσος όρος των κλασματικών κατανομών των βάσεων στους τρεις μεταφορείς μας δίνει τη σύσταση του t-RNA.

Πίνακας 5.7 Αριθμός βάσεων στον κάθε μεταφορέα

	Αριθμός βάσεων στον κάθε μεταφορέα		
Αζωτούχες βάσεις	Προλίνη-tRNA	Αργινίνη-tRNA	Θρεονίνη-tRNA
A	11	12	18
G	25	25	22
U	22	20	20
C	18	20	17
Ολικός αριθμός βάσεων	76	77	77

Πίνακας 5.8 Κλασματική κατανομή κάθε βάσης στον κάθε μεταφορέα

	Κλασματική κατανομή βάσης		
Αζωτούχες βάσεις	Προλίνη-tRNA	Αργινίνη-tRNA	Θρεονίνη-tRNA
A	0,145	0,156	0,234
G	0,329	0,325	0,286
U	0,289	0,260	0,260
C	0,237	0,260	0,221

Τελικά καταλήγουμε στα εξής αποτελέσματα για κάθε ριβονουκλεοτίδιο:

Πίνακας 5.9 Κατανομή των ριβονουκλεοτιδίων σε mRNA, rRNA, tRNA και MB κάθε ριβονουκλεοτιδίου

	mol/mol RNA			MB [g/mol RNA]
	mRNA	rRNA	tRNA	
	0,05	0,8	0,15	
AMP	0,329	0,265	0,178	329,2
GMP	0,341	0,266	0,313	345,2
UMP	0,165	0,268	0,270	306,2
CMP	0,165	0,200	0,239	305,2

Στη συνέχεια υπολογίζουμε το μέσο όρο του μοριακού κλάσματος του κάθε ριβονουκλεοτιδίου στο συνολικό RNA και πολλαπλασιάζουμε με το μοριακό του βάρος τα g ριβονουκλεοτιδίου ανά mole συνολικού RNA. Αθροίζοντας για όλα τα ριβονουκλεοτίδια βρίσκουμε ένα μέσο MB του RNA σε g/mole. Έχοντας το μέσο MB του RNA, διαιρούμε με αυτό τα μοριακά κλάσματα των ριβονουκλεοτιδίων και πολλαπλασιάζουμε με το κλάσμα βάρους του RNA στη συνολική βιομάζα (g RNA/ g ξ.μ.κ.) και με κατάλληλες μετατροπές μονάδων καταλήγουμε σε mmol νουκλεοτιδίου/ g ξ.μ.κ.. Οι μετατροπές αυτές και οι τελικοί στοιχειομετρικοί συντελεστές φαίνονται στον Πίνακα 5.10.

Πίνακας 5.10 Υπολογισμοί κλασματικής κατανομής ριβονουκλεοτιδίων στη βιομάζα σε mmol/g ξ.μ.κ.

mol/mol RNA	g νουκλ/mole RNA	g νουκλ/g RNA	mmol νουκλ/g ξ.μ.κ
0,255	84,109	0,261	0,014
0,277	95,630	0,296	0,015
0,263	80,534	0,250	0,014
0,204	62,403	0,193	0,011
Μέσο MB RNA [g/mole]	322,676		

23| Προσδιορισμός λιπιδικού περιεχομένου.

Για τον προσδιορισμό της σύστασης σε λιπαρά οξέα – κατά συνέπεια το λιπιδικό περιεχόμενο- βρέθηκαν δεδομένα βιβλιογραφικά. Η σύσταση των λιπαρών οξέων του *Nannochloropsis gaditana* προσεγγίστηκε από δεδομένα για το *Nannochloropsis occulata*, συγγενικό είδος του *N. gaditana* και φυλογενετικά όμοιο σε ποσοστό μεγαλύτερο του 95%. (Dunstan, et al., 1993).

Η σύσταση δίνεται στον Πίνακα 5.11.

Πίνακας 5.11 Σύσταση σε λιπαρά οξέα για το *Nannochloropsis* (Dunstan, et al., 1993)

Λιπαρό οξύ	% σε Ολικά λιπίδια [% mol ΛΟ/mol ΟΛ]
Μυριστικό οξύ 14:0	5
i15:0	0,3
Πενταδεκανοϊκό οξύ C15:0	0,3
Παλμιτικό οξύ C16:0	16,4
Παλμιτολεϊκό οξύ C16:1 (n-7)	21,8
C16:1 (n-5)	0,1
C16:1 (n-13) t	1
C16:2 (n-7)	1
C16:2 (n-4)	0,5
C17:1 (n-8)	0,2
Στεατικό οξύ C18:0	0,3
Ολεϊκό οξύ C18:1(n-9)	2,1
Δαμαλικό οξύ 18:1 (n-7)	0,6
Λινολεϊκό οξύ C18:2 (n-6)	1,7
α-λινολενικό οξύ C18:3 (n-3)	0,1
γ-λινολενικό οξύ C18:3 (n-6)	0,9
Αραχιδονικό οξύ C20:4 (n-6)	6,1
Εικοσαπεντανοϊκό οξύ C20:5 (n-3)	39,8
Ολικό %	98,2

Γνωρίζοντας το μοριακό βάρος κάθε λιπαρού οξέος (ΛΟ) βρίσκουμε τα γραμμάρια λιπαρού οξέος ανά mole ολικών λιπιδίων(ΟΛ). Χρησιμοποιώντας το μέσο μοριακό βάρος κάθε λιπαρού οξέος αθροίζουμε τις συνεισφορές του καθενός για να υπολογίσουμε το μέσο μοριακό βάρος μιας αλυσίδας λιπαρού οξέος. Για να συνυπολογιστεί το μοριακό βάρος του ακυλίου υπολογίστηκε ένα μέσο μοριακό βάρος λιπαρού οξέος MB(Λιπαρά οξέα χωρίς ομάδα – COOH) = 217,483 g/mol.

Πίνακας 5.1 Υπολογισμός μέσου μοριακού βάρους μιας αλυσίδας λιπαρού οξέος

Λιπαρό οξύ	MB [g ΛΟ/mole ΛΟ]	g ΛΟ/mole ΟΛ
Μυριστικό οξύ 14:0	228,4	11,420
i15:0	242,4	0,727
Πενταδεκανοϊκό οξύ C15:0	242,4	0,727
Παλμιτικό οξύ C16:0	256,4	42,050
Παλμιτολεϊκό οξύ C16:1 (n-7)	254,4	55,459
C16:1 (n-5)	254,408	0,254
C16:1 (n-13) t	254,408	2,544
C16:2 (n-7)	252	2,524
C16:2 (n-4)	252	1,260
C17:1 (n-8)	268	0,536
Στεατικό οξύ C18:0	284,48	0,853
Ολεϊκό οξύ C18:1(n-9)	282,5	5,933
Δαμαλικό οξύ 18:1 (n-7)	282,461	1,695
Λινολεϊκό οξύ C18:2 (n-6)	280,5	4,769
α-λινολενικό οξύ C18:3 (n-3)	278,43	0,278
γ-λινολενικό οξύ C18:3 (n-6)	278,43	2,506
Αραχιδονικό οξύ C20:4 (n-6)	304,47	18,573
Εικοσαπεντανοϊκό οξύ C20:5 (n-3)	302,45	120,375
Μέσο MB ενός ΛΟ [g ΛΟ/mole ΛΟ]	MB(Λιπαρά οξέα χωρίς ομάδα – COOH) [g ΛΟ/mol ΛΟ]	
272,483	217,483	

Η σύσταση για κάθε λιπαρό οξύ δίνεται σε mole ΛΟ/mole ΟΛ. Συνεπώς για να υπολογίσουμε το επιθυμητό mmole ΛΟ/ g ξ.μ.κ. χρειάζεται να υπολογιστεί το μέσο μοριακό βάρος ενός λιπιδίου. Χρειαζόμαστε το περιεχόμενο σε τριγλυκερίδια(TAGs), φωσφολιπίδια (PG, PC ,PE , PI) γαλακτολιπίδια (MGDG, DGDG), σουλφολιπίδια (SQDG) και DGDTs (betaine lipids), το οποίο βρέθηκε βιβλιογραφικά. (Schneider & Roessler, 1994)

Οι υπολογισμοί αυτοί δίνονται στον Πίνακα 5.13.

Πίνακας 5.13 Μολαρική σύσταση λιπιδίων και μοριακά βάρη (Schneider & Roessler, 1994) (James Hutton Institute, 2014)

	Αριθμός ΛΟ που περιέχει κάθε λιπίδιο	ΜΒ βασικής δομής(ή Ζ για γλυκεροφωσfolιπίδια)	ΜΒ λιπιδίου	% mole λιπιδίου/mole ΟΛ
Τριγλυκερίδια (TAG)	3	89	741,449	50,4
Φωσfolιπίδια				
φωσφατιδυλογλυκερόλη (PG)	2	75	733,936	3,8
Φωσφατιδυλοχολίνη (PC)	2	90	748,936	11,6
Φωσφατιδυλοαιθανολαμίνη (PE)	2	48	706,936	3
Φωσφατιδυλινοσιτολη (PI)	2	179	837,936	2
Γλυκοζυλογλυκερίδια (Γαλακτολιπίδια & Σουλfolιπίδια)				
MGDG	2	308	742,966	11,2
DGDG	2	470	904,966	8,8
SQDG	2	371	805,966	3,4
Βεταϊνικά λιπίδια				
DGDTS	2	292	726,966	5,8

Το μοριακό βάρος κάθε λιπιδίου υπολογίζεται ως εξής:

- Για τα τριγλυκερίδια αθροίζουμε το ΜΒ της βασικής δομής με το γινόμενο του αριθμού λιπαρών οξέων που περιέχουν με το μέσο μοριακό βάρος λιπαρού οξέος που βρήκαμε πριν (χωρίς την ομάδα –COOH), δηλαδή $89 + 3 \cdot 217,483 = 741,449$ g/mole.
- Για τα φωσfolιπίδια αθροίζουμε το μοριακό βάρος της βασικής δομής χωρίς τα λιπαρά οξέα ή το Ζ (η τιμή αυτού του μοριακού βάρους για γλυκεροφωσfolιπίδια βρέθηκε ίση με 223, 87 g/mole) με το γινόμενο του αριθμού λιπαρών οξέων που περιέχουν με το μέσο μοριακό βάρος λιπαρού οξέος που βρήκαμε πριν (χωρίς την ομάδα –COOH) και με το μοριακό βάρος βασικής δομής-Z- (όπως δίνεται στην 3^η στήλη του παραπάνω πίνακα), δηλαδή για το PG: $223,97 + 2 \cdot 217,483 + 75 = 733,936$ g/mole.
- Για τα γλυκοζυλογλυκερίδια και τα βεταϊνικά λιπίδια αθροίζουμε το ΜΒ της βασικής δομής με το γινόμενο του αριθμού λιπαρών οξέων που περιέχουν με το μέσο μοριακό βάρος λιπαρού οξέος που βρήκαμε πριν (χωρίς την ομάδα –COOH), δηλαδή για το MGDG: $308 + 2 \cdot 217,483 = 742,966$ g/mol.

Έχοντας το μοριακό βάρος κάθε λιπιδίου καθώς και το μολαρικό του κλάσμα στα συνολικά λιπίδια το γινόμενο τους μας δίνει τη συνεισφορά του κάθε λιπιδίου, αθροίζοντας τις συνεισφορές βρίσκουμε το μέσο μοριακό βάρος ενός λιπιδίου, το οποίο είναι $MB_{\text{λιπιδίου}}=758,839 \text{ g/mole}$.

Διαιρώντας το μολαρικό κλάσμα κάθε λιπαρού οξέος με το μέσο μοριακό βάρος λιπιδίου βρίσκουμε τα mole λιπαρού οξέος ανά γραμμάριο λιπιδίου. Γνωρίζοντας την κατά βάρος σύσταση σε ξηρή μάζα κυττάρων των λιπιδίων και με κατάλληλες μετατροπές μονάδων βρίσκουμε τελικά για κάθε λιπαρό οξύ τα mmol λιπαρού οξέος/ g ξ.μ.κ.. Οι τελικοί στοιχειομετρικοί συντελεστές δίνονται στον Πίνακα 5.14.

Πίνακας 5.14 Κατανομή λιπαρών οξέων στη βιομάζα σε mmol/g ξ.μ.κ.

Λιπαρό οξύ	mole ΛΟ/g ΟΛ	mmol ΛΟ/g ξ.μ.κ.
Μυριστικό οξύ 14:0	6,61263E-05	0,0265
i15:0	3,96758E-06	0,0016
Πενταδεκανοϊκό οξύ C15:0	3,96758E-06	0,0016
Παλμιτικό οξύ C16:0	0,000216894	0,0868
Παλμιτολεϊκό οξύ C16:1 (n-7)	0,000288311	0,1153
C16:1 (n-5)	1,32253E-06	0,0005
C16:1 (n-13) t	1,32253E-05	0,0053
C16:2 (n-7)	1,32253E-05	0,0053
C16:2 (n-4)	6,61263E-06	0,0026
C17:1 (n-8)	2,64505E-06	0,0011
Στεατικό οξύ C18:0	3,96758E-06	0,0016
Ολεϊκό οξύ C18:1(n-9)	2,7773E-05	0,0111
Δαμαλικό οξύ 18:1 (n-7)	7,93515E-06	0,003
Λινολεϊκό οξύ C18:2 (n-6)	2,24829E-05	0,0090
α-λινολενικό οξύ C18:3 (n-3)	1,32253E-06	0,0005
γ-λινολενικό οξύ C18:3 (n-6)	1,19027E-05	0,0048
Αραχιδονικό οξύ C20:4 (n-6)	8,06741E-05	0,0323
Εικοσαπεντανοϊκό οξύ C20:5 (n-3)	0,000526365	0,2105

Από αυτά τα λιπαρά οξέα στην εξίσωση της βιομάζας συμμετέχουν τα εξής: μυριστικό, παλμιτικό, παλμιτολεϊκό, στεατικό, ολεϊκό, λινολεϊκό, α-λινολενικό, γ-λινολενικό, αραχιδονικό και το εικοσαπεντανοϊκό.

24| Προσδιορισμός σύστασης σακχάρων(υδατανθράκων)

Για τον προσδιορισμό της σύστασης σε σάκχαρα βρέθηκαν δεδομένα βιβλιογραφικά. Η σύσταση των σακχάρων του *Nannochloropsis gaditana* προσεγγίστηκε από δεδομένα για το *Nannochloropsis oculata*, συγγενικό είδος του *N. gaditana* και φυλογενετικά όμοιο σε ποσοστό μεγαλύτερο του 95%. (Brown, 1991)

Πίνακας 5.15 Σύσταση σε σάκχαρα του *Nannochloropsis* (Brown, 1991)

Σάκχαρο	%κ.β. [% g σακχάρου/g πολυσακχαριτών]	g σακχάρου/g πολυσακχαριτών]
Γλυκόζη	68,2	0,682
Ραμνόζη	8,3	0,083
Μαννόζη	6,1	0,061
Ριβόζη	4,6	0,046
Ξυλόζη	4,4	0,044
Φουκόζη	4,4	0,044
Γαλακτόζη	3,8	0,038

Με χρήση του ποσοστού κατά βάρος καθώς και του μοριακού βάρους κατάλοιπου σακχάρου (αφού ο πολυμερισμός συνήθως σχετίζεται με την απώλεια ενός μορίου νερού) υπολογίζουμε τα mole σακχάρου ανά γραμμάριο πολυσακχαριτών. Γνωρίζοντας τη σύσταση των υδατανθράκων σε πολυσακχαρίτες (0,88 g πολυσακχαριτών/ g υδατανθράκων) αλλά και τη σύσταση της βιομάζας σε υδατάνθρακες (0,078 g υδατανθράκων/ g ξ.μ.κ.) και με κατάλληλες μετατροπές μονάδων καταλήγουμε σε mmol σακχάρου/g ξ.μ.κ, που θα αποτελέσει το στοιχειομετρικό συντελεστή του κάθε σακχάρου στην συνολική αντίδραση-εξίσωση της βιομάζας. Παρακάτω δίνεται ο Πίνακας 5.16 των μετατροπών και τελικών στοιχειομετρικών συντελεστών.

Πίνακας 5.16 Υπολογισμοί κλασματικής κατανομής σακχάρων στη βιομάζα σε mmol/g ξ.μ.κ.

Σάκχαρο	MB καταλοίπου [g σακχάρου/mole σακχάρου]	mol σακχάρου/g πολυσακχαρίτη	mmol σακχάρου/g ξ.μ.κ.
Γλυκόζη	162	0,004210	0,289
Ραμνόζη	145	0,000572	0,039
Μαννόζη	162	0,000377	0,026
Ριβόζη	132	0,000348	0,024
Ξυλόζη	132	0,000333	0,023
Φουκόζη	146	0,000301	0,021
Γαλακτόζη	162	0,000235	0,016

Από αυτά τα σάκχαρα στην εξίσωση της βιομάζας συμμετέχουν τα εξής: Γλυκόζη, μαννόζη, ριβόζη, ξυλόζη, γαλακτόζη και άμυλο το οποίο θεωρήθηκε ότι συνεισφέρει κατά το ήμισυ της γλυκόζης στη βιομάζα, λόγω έλλειψης δεδομένων.

25| Προσδιορισμός περιεχομένου σε πολυαμίνες και βιταμίνες.

Το περιεχόμενο σε πολυαμίνες –συγκεκριμένα πουτρεσίνη και σπερμιδίνη- βρέθηκε βιβλιογραφικά και συγκεκριμένα για το στέλεχος *Nannochloropsis occulata*, συγγενικό είδος του *Nannochloropsis gaditana*. (Hamana, et al., 2004)

Πίνακας 5.17 Σύσταση σε πολυαμίνες για το *Nannochloropsis* (Hamana, et al., 2004)

Πολυαμίνη	μmol/g υγρό βάρος κυττάρου
Πουτρεσίνη	0,65
Σπερμιδίνη	1,32

Για να βρούμε τα mmol πολυαμίνης/g ξ.μ.κ. χρησιμοποιούμε την αναλογία υγρής μάζας/ ξηρής μάζας, η οποία βρέθηκε βιβλιογραφικά και είναι 9,577 (Augustine, et al., 2014). Τα αποτελέσματα φαίνονται Πίνακα 5.18.

Πίνακας 5.18 Κλασματική κατανομή πολυαμινών στη βιομάζα σε mmol/g ξηρής μάζας κυττάρων

Πολυαμίνη	mmol/g υ.μ.κ	mmol/g ξ.μ.κ.
Πουτρεσίνη	0,00065	0,006
Σπερμιδίνη	0,00132	0,013

Τα μικροφύκη αποτελούν πλούσια πηγή δύο βασικών βιταμινών, του ασκορβικού οξέος (βιταμίνη C) και της ριβοφλαβίνης. Η σύσταση αυτών των δύο βιταμινών στο *Nannochloropsis* βρέθηκε βιβλιογραφικά. (Brown, et al., 1997) και δίνεται στον Πίνακα 5.19.

Πίνακας 5.19 Σύσταση σε βιταμίνες για το *Nannochloropsis* (Brown, et al., 1997) και μοριακά βάρη

Βιταμίνη	mg/g ξ.μ.κ.	MB [g/mol]
Ασκορβικό οξύ	8	176,12
	μg/g ξ.μ.κ.	MB [g/mol]
Ριβοφλαβίνη	38	376,6

Διαιρώντας τη σύσταση με το μοριακό βάρος και με κατάλληλες μετατροπές μονάδων βρίσκουμε τα mmol βιταμίνης/g ξ.μ.κ..

Πίνακας 5.20 Κλασματική κατανομή βιταμινών στη βιομάζα σε mmol/g ξηρής μάζας κυττάρων

Βιταμίνη	mmol/g ξ.μ.κ.
Ασκορβικό οξύ	0,045
Ριβοφλαβίνη	0,0001

26| Προσδιορισμός σύστασης σε χλωροφύλλη α και καροτενοειδή.

Το *Nannochloropsis* είναι ένα ξεχωριστό φύκος διότι δεν παράγει καμία άλλη χλωροφύλλη εκτός της χλωροφύλλης α. Η σύστασή της βρέθηκε βιβλιογραφικά. (Reboloso-Fuentes, et al., 2001)

Πίνακας 5.21 Σύσταση σε χλωροφύλλη για το *Nannochloropsis* (Reboloso-Fuentes, et al., 2001)

%κ.β. [% g χλωροφύλλης/ g ξ.μ.κ.]

Χλωροφύλλη α	6
---------------------	----------

Διαιρώντας τη σύσταση με το μοριακό βάρος της χλωροφύλλης και με κατάλληλες μετατροπές μονάδων καταλήγουμε στα mmol χλωροφύλλης α/g ξ.μ.κ., όπως δίνονται στον Πίνακα 5.22

Πίνακας 5.22 Κλασματική κατανομή χλωροφύλλης στη βιομάζα σε mmol/g ξ.μ.κ

	MB [g/mol]	mmol χλωροφύλλης /g ξ.μ.κ
Χλωροφύλλη α	893,49	0,0672

Στον Πίνακα 5.23 δίνονται οι συστάσεις καροτενοειδών, οι οποίες βρέθηκαν βιβλιογραφικά. (Lubián, et al., 2000)

Πίνακας 5.23 Σύσταση σε βιταμίνες για το *Nannochloropsis* (Lubián, et al., 2000)

Καροτενοειδή	%κ.β. στα ολικά καροτενοειδή	%κ.β. [g καροτενοειδούς/g ξ.μ.κ.]
Ασταξανθίνη	13	0,7
Ζεαξανθίνη		0,6
Βιολαξανθίνη	60	3,96
Ανθεραξανθίνη	[21% της βιολαξανθίνης μετατρέπεται σε ανθεραξανθίνη]	

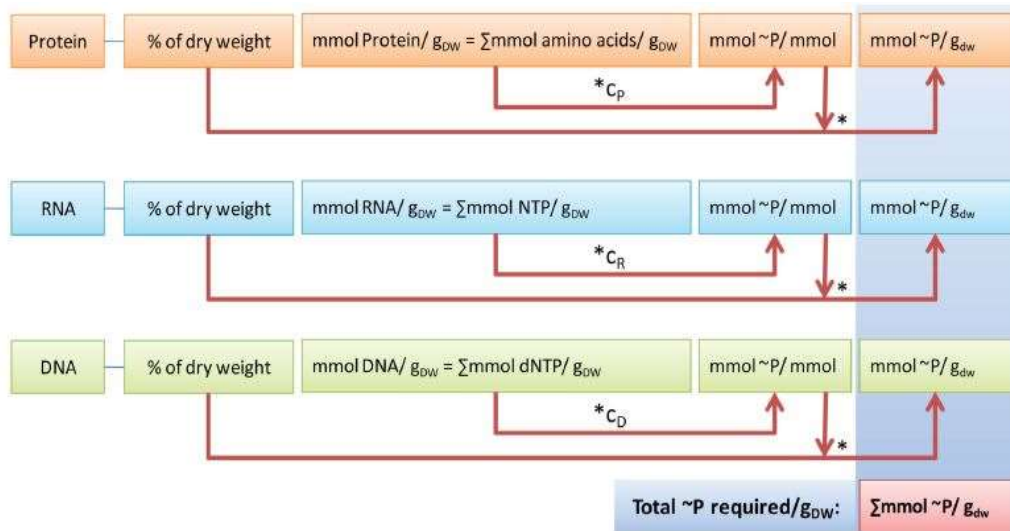
Διαιρώντας τη σύσταση με το μοριακό βάρος του κάθε καροτενοειδούς και με κατάλληλες μετατροπές μονάδων καταλήγουμε σε mmol καροτενοειδούς/g ξ.μ.κ..

Πίνακας 5.24 Κλασματική κατανομή καροτενοειδών στη βιομάζα σε mmol/g ξ.μ.κ.

Καροτενοειδή	MB [g/mol]	mmol/g ξ.μ.κ.
Ασταξανθίνη	596,84	0,0117
Ζεαξανθίνη	568,88	0,0105
Βιολαξανθίνη	600,85	0,0659
Ανθεραξανθίνη	584,87	0,0142

27| Προσδιορισμός GAM- αντίδρασης ATP συντήρησης που σχετίζεται με την ανάπτυξη.

Η GAM αντιπροσωπεύει την ενέργεια (σε μορφή ATP) που είναι απαραίτητη για την αναπαραγωγή του κυττάρου, συμπεριλαμβάνοντας τη μακρομοριακή σύνθεση (π.χ. πρωτεΐνες, DNA και RNA). Στην Εικ.10 απεικονίζεται ο τρόπος υπολογισμού της GAM χρησιμοποιώντας την ολική ποσότητα (mmol) μακρομορίων (Πρωτεΐνες, DNA και RNA) και τη γνωστή ποσότητα δεσμών φωσφορικού άλατος που είναι απαραίτητη για τη σύνθεση του μακρομορίου.



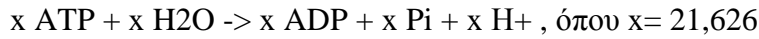
Εικόνα 5.10 Υπολογισμός κόστους σε ATP για την ανάπτυξη

Στον Πίνακα 5.25 φαίνονται οι υπολογισμοί που πραγματοποιήθηκαν.

Πίνακας 5.25 Υπολογισμοί GAM

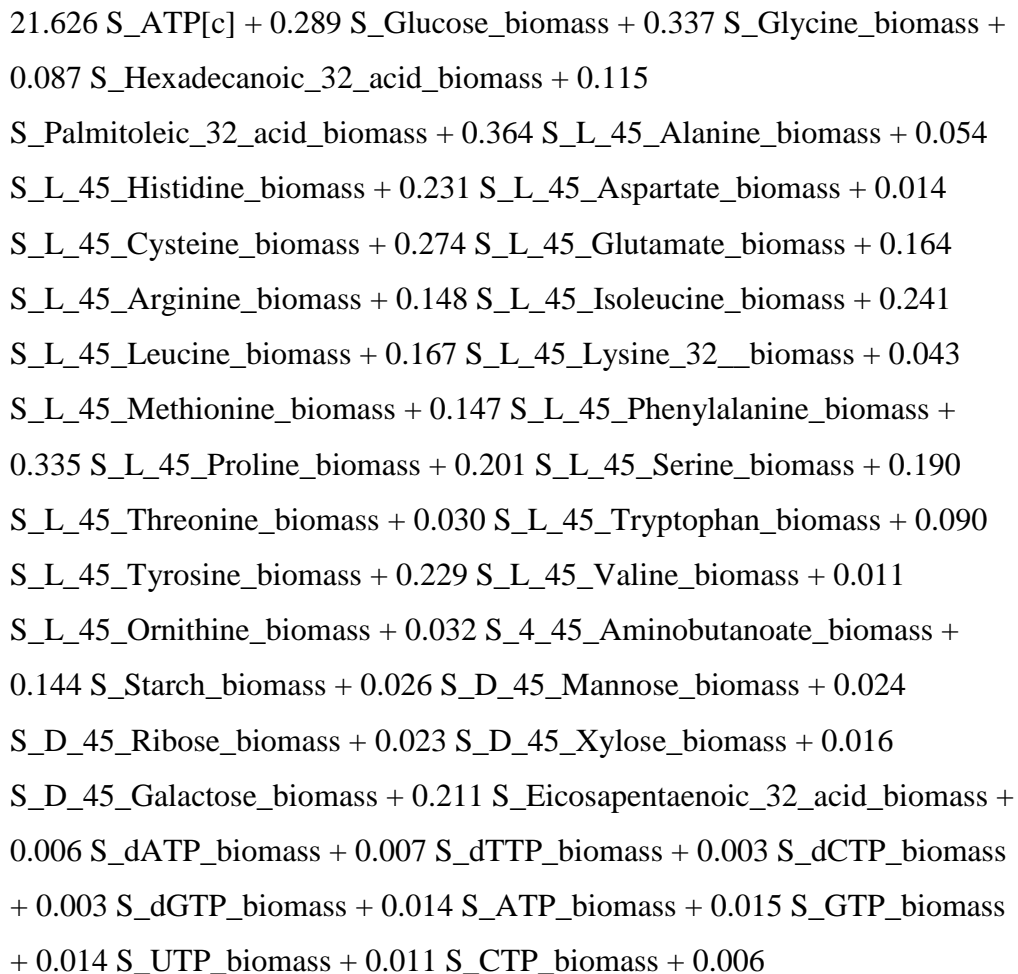
	%κ.β.	mmol/gDW	C _P ,C _D ,C _R	mmol ~P/mmol	ολικά Σmmol ~P/gDW	Ολικά
Protein	0,35	3,305	4,324	14,290	5,001	21,6261
DNA	0,0061	0,0197	1,365	0,0269	0,0001	0,000223
RNA	0,0171	0,053	0,406	0,0215	0,00036	0,000149
					Total	21,626

Η GAM μας δείχνει ότι απαιτείται υδρόλυση 21,626 mmol ATP/g ξ.μ.κ. και προστίθεται στην αντίδραση της βιομάζας:



28| Σύνταξη της αντίδρασης βιομάζας και προσθήκη στο μοντέλο.

Η αντίδραση βιομάζας που προστέθηκε στο μοντέλο για το *Nannochloropsis* είναι:



$S_Putrescine_biomass + 0.013 S_Spermidine_biomass + 0.01$
 $S_Riboflavin_biomass + 0.045 S_Ascorbate_biomass + 0.011$
 $S_40_9Z_41_45_Octadecenoic_32_acid_biomass + 0.026$
 $S_Tetradecanoic_32_acid_biomass + 0.002$
 $S_Octadecanoic_32_acid_biomass + 0.009$
 $S_40_9Z_44_12Z_41_45_Octadecadienoic_32_acid_biomass + 0.001$
 $S_40_9Z_44_12Z_44_15Z_41_45_Octadecatrienoic_32_acid_biomass +$
 $0.005 S_40_6Z_44_9Z_44_12Z_41_45_Octadecatrienoic_32_acid_biomass$
 $+ 0.032$
 $S_40_5Z_44_8Z_44_11Z_44_14Z_41_45_Icosatetraenoic_32_acid_biomas$
 $s + 0.067 S_Chlorophyll_32_a_biomass + 0.0659 S_Violaxanthin_biomass +$
 $0.0105 S_Zeaxanthin_biomass + 0.0117 S_Astaxanthin_biomass + 0.0142$
 $S_Antheraxanthin_biomass \rightarrow 21.626 S_ADP[c] + 21.626$
 $S_Orthophosphate[c]$

29| Προσθήκη αντιδράσεων ζήτησης (demand reactions).

Οι αντιδράσεις ζήτησης είναι ασύμμετρες αντιδράσεις που επιτρέπουν τη συσσώρευση μίας ένωσης, κάτι το οποίο δεν επιτρέπεται σε μοντέλα steady-state λόγω των απαιτήσεων εξισορρόπησης της μάζας. Προστέθηκαν αντιδράσεις ζήτησης για όλες εκείνες τις ενώσεις που παράγονται από τον οργανισμό και συμμετέχουν στην αντίδραση της βιομάζας.

30| Προσθήκη αντιδράσεων βύθισης (sink reactions).

Οι αντιδράσεις αυτές είναι παρόμοιες με τις ζήτησης αλλά ορίζονται ως αντιστρεπτές. Χρησιμοποιούνται για ενώσεις που παράγονται από μη μεταβολικές κυτταρικές διεργασίες αλλά πρέπει να μεταβολιστούν και συμμετέχουν στο δίκτυο. Πρέπει να προστίθενται με προσοχή καθώς υπερβολικός αριθμός τέτοιων αντιδράσεων καθιστά ικανό το μοντέλο να αναπτύσσεται χωρίς να υπάρχουν πόροι στο μέσο.

31| Προσδιορισμός απαιτήσεων για ανάπτυξη.

Κατασκευή λίστας πληροφοριών για το μέσο ανάπτυξης και τη σύσταση αυτού. Αυτές οι πληροφορίες αποτελούν σημαντική βοήθεια για τα επόμενα δύο στάδια.

3^ο ΣΤΑΔΙΟ : ΜΕΤΑΤΡΟΠΗ ΤΟΥ ΠΡΟΣΧΕΔΙΟΥ ΣΕ ΜΑΘΗΜΑΤΙΚΟ ΜΟΝΤΕΛΟ (Conversion from reconstruction to mathematical model)

32| Έναρξη του COBRA Toolbox.

Αφού έχει γίνει εγκατάσταση του Matlab και των απαραίτητων «εργαλειοθηκών» (SBML Toolbox και COBRA Toolbox) και ενός LP solver γίνεται έναρξη του COBRA Toolbox με την εισαγωγή της εντολής *initCobraToolbox* στη γραμμή εντολών του Matlab.

33| Φόρτιση κατασκευής στο Matlab.

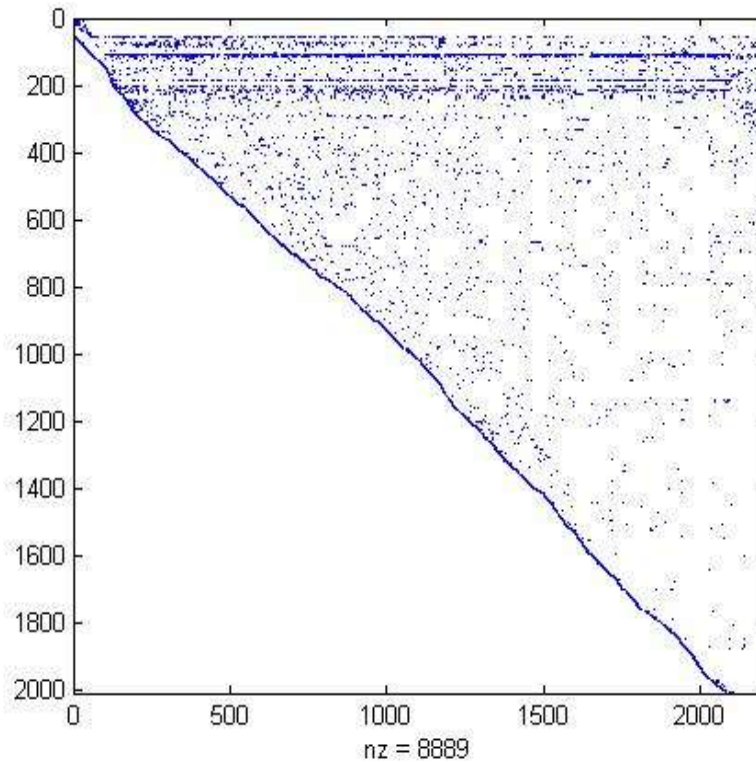
Το μοντέλο-κατασκευή έχει φτιαχτεί ως τώρα σε ένα αρχείο excel. Το πρώτο φύλλο σε αυτό το αρχείο είναι η λίστα των αντιδράσεων, όπως έχει υποδειχθεί στην Εικ.2, και έχει ονομασία “reactions”. Το δεύτερο φύλλο στο αρχείο excel είναι η λίστα των μεταβολιτών, όπως έχει υποδειχθεί στην Εικ.3 και έχει ονομασία “metabolites”. Χρησιμοποιείται η παρακάτω εντολή για να εισαχθεί το προσχέδιο στο Matlab και να μετατραπεί σε μαθηματικό μοντέλο:

```
model=xls2model('fileName'), όπου filename τό όνομα του αρχείου excel, π.χ. model.xls
```

Το μεταβολικό μοντέλο που εισάγαμε αποθηκεύεται σε μία δομή με ονομασία “model” στο Matlab. Αυτή η κατασκευή περιέχει όλες τις πληροφορίες για το μοντέλο.

34| Επιβεβαίωση πίνακα S.

Με την εντολή *spy(model.S)* γίνεται επιβεβαίωση της δομής του εισηγμένου πίνακα S. Αυτή η απεικόνιση θα πρέπει να επαναλαμβάνεται όταν προστίθενται ή αφαιρούνται αντιδράσεις στο μοντέλο για να εξασφαλίζεται ότι είναι συνδεδεμένες στο δίκτυο.



Εικόνα 5.19 Πίνακας-δίκτυο S για το μοντέλο του *Nannochloropsis gaditana*

35| Ορισμός αντικειμενικής συνάρτησης.

Με την εντολή `model=changeObjective(model,rxnNameList)`, όπου `rxnNameList` δίνουμε το όνομα της αντίδρασης(-εων) που θέλουμε να γίνει αντικειμενική συνάρτηση και δίνεται αυτόματα ο συντελεστής 1 (`objectiveCoefficient`).

36| Επιβολή περιορισμών προσομείωσης.

Με την εντολή `model=changeRxnBounds(model,rxnNameList,value,boundType)`, όπου `value`: ο αριθμός-όριο, `boundType`: 'b'-και τα δύο όρια ίδια τιμή; 'l'-κάτω όριο; 'u'-άνω όριο, ορίζονται οι περιορισμοί της προσομείωσης.

Με την εντολή `printConstraints(model,MinInf,MaxInf)`, λαμβάνουμε τη λίστα των περιορισμένων αντιδράσεων που είναι μεγαλύτερες μιας ελάχιστης τιμής και μικρότερες μιας μέγιστης τιμής.

Επιπλέον υπάρχει η εντολή `printFluxVector(model,fluxData)` που παρέχει τη λίστα όλων των αντιδράσεων με τις ροές τους σε μια λύση.

4^ο ΣΤΑΔΙΟ : ΑΞΙΟΛΟΓΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ (Network evaluation='Debugging mode')

Δοκιμή αν το μοντέλο είναι ισορροπημένο ως προς τη μάζα και τα φορτία

37| Έλεγχος για μη ισορροπημένες στοιχειομετρικά αντιδράσεις.

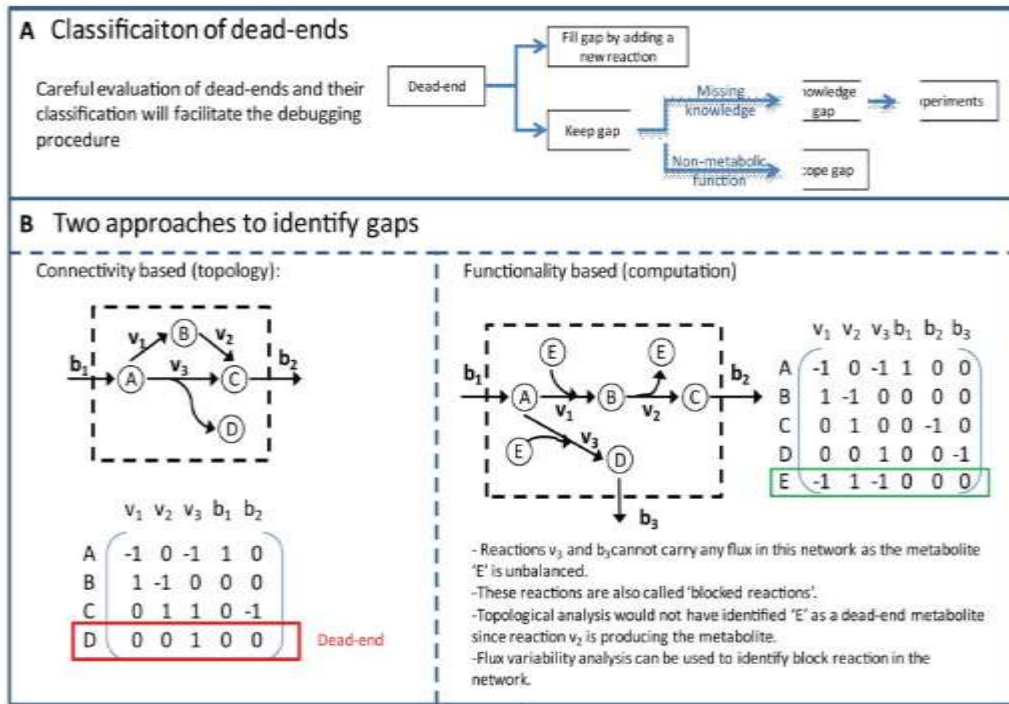
Με την εντολή $[UnbalancedRxns]=checkMassChargeBalance(model,RxnList)$, επιστρέφεται μία λίστα με τις αντιδράσεις που δεν είναι στοιχειομετρικά ισορροπημένες και ποια στοιχεία αυτών συγκεκριμένα δεν έχουν τη σωστή στοιχειομετρία.

38| Αξιολόγηση των μη ισορροπημένων στοιχειομετρικά αντιδράσεων.

Κάποιες αντιδράσεις είναι πάντα μη ισορροπημένες, όπως η αντίδραση της βιομάζας, οι αντιδράσεις ζήτησης, βύθισης αλλά και ανταλλαγής. Για τις υπόλοιπες ελέγχεται αν λείπει κάποιο πρωτόνιο ή/και μόριο νερού ή αν ο στοιχειομετρικός συντελεστής κάποιου μεταβολίτη είναι λάθος. Αν λείπει κάποιο πρωτόνιο ως υπόστρωμα, ένας δότης πρωτονίων μπορεί να είναι απαραίτητος (π.χ. NADH, NADPH). Αυτό απαιτεί χρήση βιβλιογραφίας για να αναγνωριστεί ο κατάλληλος δότης. Αν λείπει μόριο νερού, μετά την προσθήκη του θα χρειαστεί να εξισορροπηθούν το πρωτόνιο και το οξυγόνο πάλι.

39| Αναγνώριση αδιέξοδων (dead-end) μεταβολιτών.

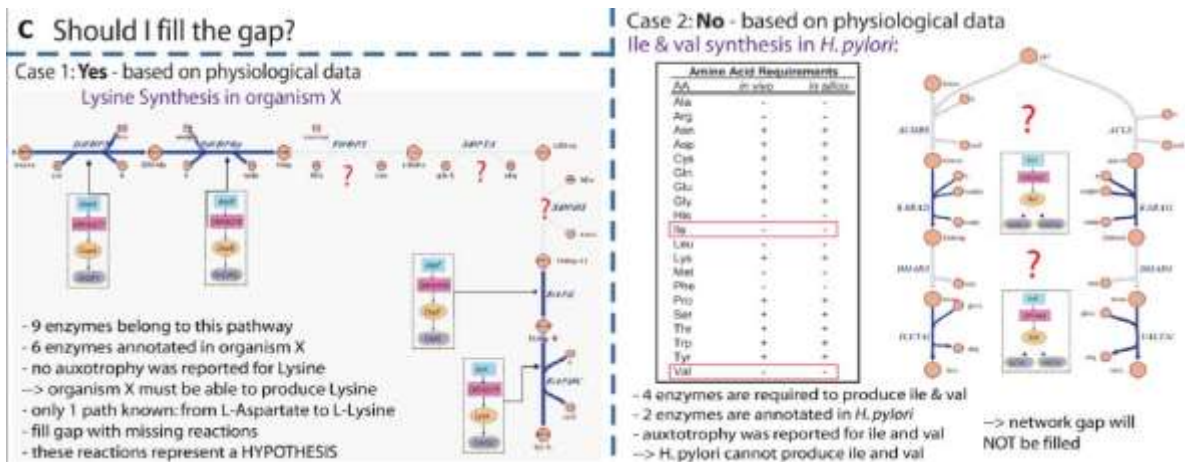
Με την εντολή $[allGaps,rootGaps,downstreamGaps] = gapFind(model)$, εντοπίζονται τα κενά στο μοντέλο. Η εντολή αυτή επιστρέφει μία λίστα των μεταβολιτών αυτών ('κενά', 'Gaps') που είτε μόνο παράγονται είτε μόνο καταναλώνονται στο δίκτυο. Μέχρι αυτό το σημείο έχει ολοκληρωθεί η πρώτη επανάληψη της χειροκίνητης βελτίωσης του μοντέλου. Η ανάλυση κενών περιλαμβάνει την αναγνώριση αλλά και την δοκιμαστική συμπλήρωση των κενών του δικτύου. Στις Εικόνες 13, 14 επεξηγείται η διαδικασία ανάλυσης κενών.



Εικόνα 5.12 Ανάλυση κενών

Στην Εικ. 12-A καταδεικνύεται ότι αν και πολλοί αδιέξοδοι μεταβολίτες που δημιουργούν «κενά» μπορούν να συνδεθούν με το υπόλοιπο δίκτυο με επαναξιολόγηση γονιδιακών και πειραματικών δεδομένων, αρκετοί από αυτούς θα παραμείνουν στο βελτιωμένο μοντέλο. Αυτοί οι αδιέξοδοι μεταβολίτες χωρίζονται σε δύο κατηγορίες, ανάλογα με τον τύπο της αντίδρασης που θα μπορούσε να τους συνδέσει με το υπόλοιπο δίκτυο: σε «κενά» γνώσης και σε «κενά» πεδίου. Παρόλο που τα «κενά» γνώσης αντιπροσωπεύουν έλλειψη βιοχημικής γνώσης για τον συγκεκριμένο οργανισμό, τα «κενά» δικτύου περιλαμβάνουν αντιδράσεις και κυτταρικές διεργασίες, οι οποίες δεν λογαριάζονται στην κατασκευή μεταβολικών δικτύων (π.χ. η μεθυλίωση του DNA). Στην Εικ. 12-B απεικονίζονται οι δύο προσεγγίσεις αναγνώρισης κενών στο μοντέλο. Ο πρώτος τρόπος προσέγγισης αφορά τη συνδεσιμότητα, οπότε μπορεί κανείς να μετρήσει τις μη-μηδενικές καταχωρίσεις σε κάθε σειρά του πίνακα S και να ταυτοποιήσει αυτούς τους μεταβολίτες, οι οποίοι μόνο παράγονται ή μόνο καταναλώνονται. Στο παράδειγμα που φαίνεται, ο μεταβολίτης D μόνο παράγεται από την αντίδραση v_3 και ο πίνακας S περιέχει μόνο μία καταχώριση στη σειρά που να αφορά το μεταβολίτη D. Η δεύτερη προσέγγιση βασίζεται στη λειτουργικότητα του μοντέλου, δηλαδή την ικανότητα του μοντέλου να φέρει ροή διαμέσου κάθε αντίδρασης του δικτύου. Αυτό μπορεί να ελεγχθεί με

αναγνώριση «μπλοκαρισμένων» αντιδράσεων (blocked reactions), οι οποίες άμεσα ή έμμεσα συνδέονται με έναν ή περισσότερους αδιέξοδους μεταβολίτες. Στο παράδειγμα που φαίνεται, δεν θα μπορούσε να αναγνωριστεί ο μεταβολίτης E ως αδιέξοδος βάσει τις προσέγγισης συνδεσιμότητας καθώς και παράγεται και καταναλώνεται στο δίκτυο. Ωστόσο, μέσω ελέγχου ροών διαμέσου των αντιδράσεων που περιέχουν τον E αποδεικνύεται ότι η αντίδραση v_3 και b_3 δεν φέρουν καθόλου ροή.



Εικόνα 5.13 Ανάλυση κενών. Υποδεικνύονται δύο παραδείγματα που αφορούν το ερώτημα συμπλήρωσης ή μη ενός κενού.

40| Εξεύρεση υποψήφιων αντιδράσεων για τη συμπλήρωση «κενών».

Με χρήση βιβλιογραφίας, καθώς και εργαλεία γονιδιακών πληροφοριών γίνεται η εύρεση υποψηφίων αντιδράσεων για τη συμπλήρωση κάθε κενού. Συγκεκριμένα για το *Nannochloropsis* χρησιμοποιήθηκε εκτεταμένα η πύλη αναζήτησης γονιδίων του, είτε μέσω ονομασίας γονιδίου είτε μέσω του εργαλείου Blast. Δηλαδή μέσω της βάσης Kegg βρισκόταν η κατάλληλη αντίδραση για τη συμπλήρωση του κενού αλλά δεν υπήρχε γονιδιακή πληροφορία για το *Nannochloropsis*, ενώ υπήρχε για συγγενικά του είδη. Μέσω Blast γίνεται σύγκριση της πρωτεϊνικής αλληλουχίας του συγγενικού οργανισμού με το γονιδίωμα του *Nannochloropsis*. Ως αποτέλεσμα δίνεται το γονίδιο του *Nannochloropsis* το οποίο ταυτίζεται περισσότερο με εκείνο του συγγενικού οργανισμού. Συγγενικοί οργανισμοί θεωρήθηκαν το μικροφύκος

Chlamydomonas reinhardtii, το *Araridopsis thaliana* καθώς και κάποια είδη κόκκινων και πράσινων φυκών. (The CRIBI Biotechnology Center, 2013)

The image shows a web interface for BLASTing against the *N. gaditana* genome. At the top, there is a blue header with the BLAST logo and the text "N. gaditana genome". Below this is a large text input area labeled "Nucleotide sequence or sequences (FASTA, multi FASTA or plain sequence)". To the left of this area is a "paste sample" button. Below the input area, there are two sections: "Program:" and "Database:". The "Program:" section has five radio buttons: "BlastN" (which is selected), "BlastP", "BlastX", "tBlastN", and "tBlastX". The "Database:" section has five radio buttons: "Genome" (which is selected), "Genes", "Proteins", "Chl only", and "Mito only".

Εικόνα 5.14 Εργαλείο σύγκρισης πρωτεϊνικών αλληλουχιών με το γονιδίωμα του *Nannochloropsis gaditana* (The CRIBI Biotechnology Center, 2013)

41| Προσθήκη αντιδράσεων συμπλήρωσης κενών στο μοντέλο.

Εφόσον πειραματικά ή γονιδιωματικά δεδομένα υποστηρίζουν αυτές τις αντιδράσεις ή είναι απαραίτητες για λόγους μοντελοποίησης προστίθενται στο υπόλοιπο δίκτυο, και ακολουθούνται τα βήματα 5 έως 11. Δίνεται μεγάλη προσοχή στο να μη δημιουργηθούν καινούρια «κενά» στο δίκτυο με την προσθήκη αυτών των αντιδράσεων, δηλαδή οι μεταβολίτες που περιέχουν να είναι συνδεδεμένοι με το υπόλοιπο δίκτυο. Επαναλαμβάνεται το βήμα 41 για επιβεβαίωση.

Η διαδικασία της ανάλυσης κενών μπορεί να έχει ως αποτέλεσμα τη συμπλήρωση αντιδράσεων μεταφοράς και ανταλλαγής. Επαναλαμβάνεται το βήμα 18.

42| Ορισμός περιορισμών ανταλλαγής για μία κατάσταση προσομοίωσης.

Προσδιορίζεται μία περιβαλλοντική κατάσταση στην οποία θα πραγματοποιηθούν αρχικά οι περισσότεροι έλεγχοι αξιολόγησης του δικτύου ('standard condition'). Με την εντολή *model=changeRxnBounds(model, rxnNameList,value,boundType)*, όπου *value*: ο αριθμός-όριο, *boundType*: 'b'- και τα δύο όρια ίδια τιμή; 'l'-κάτω όριο; 'u'-άνω όριο, ορίζονται οι περιορισμοί της προσομείωσης.

Δοκιμή αν οι πρόδρομες ουσίες της βιομάζας (biomass precursors) μπορούν να παραχθούν σε πρότυπο μέσο (standard medium)

Κατά την κατασκευή του προσχεδίου έχουν προστεθεί για το *Nannochloropsis* οι αντιδράσεις ζήτησης (demand reactions) για κάθε συστατικό της βιομάζας.

Για κάθε συστατικό i της βιομάζας πραγματοποιείται ο ακόλουθος έλεγχος:

43| Ορισμός της αντίδρασης ζήτησης (του συστατικού i) ως αντικειμενική συνάρτηση.

model=changeObjective(model,rxnName);

44| Μεγιστοποίηση ('max') για την καινούρια αντικειμενική συνάρτηση (Demand function).

FBAsolution = optimizeCbModel(model,'max')

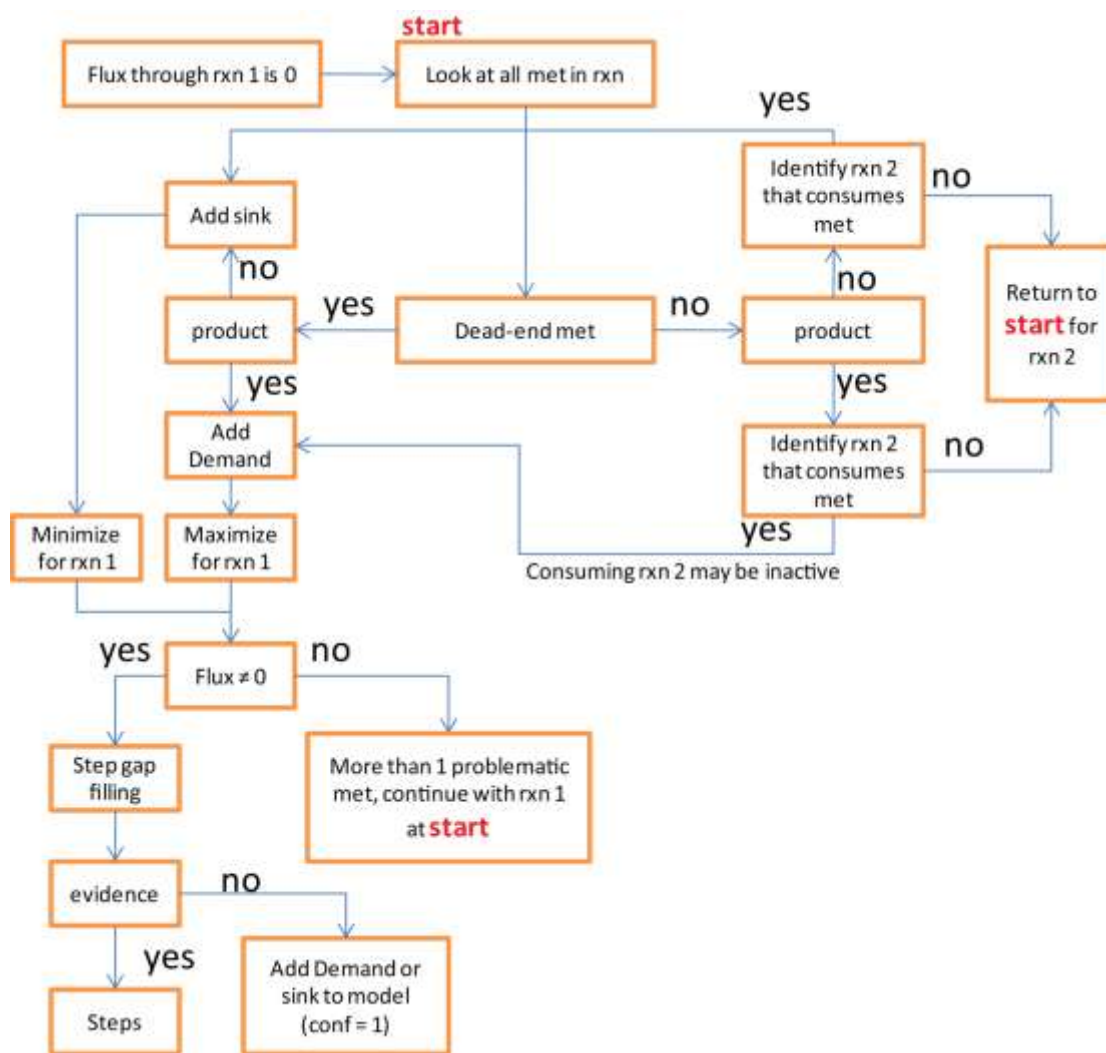
Η δομή 'FBAsolution' περιέχει την βέλτιστη λύση διανύσματος ('FBAsolution.x') και την τιμή της αντικειμενικής συνάρτησης ('FBAsolution.obj'). Αν είναι *FBAsolution.obj>0* σημαίνει ότι το μοντέλο μπορεί να παράξει το συστατικό αυτό και προχωράμε με το επόμενο συστατικό i της βιομάζας. Αν είναι *FBAsolution.obj =0* το μοντέλο δεν μπορεί να παράγει το συστατικό της βιομάζας και ακολουθούνται τα βήματα 45 και 46 για την επίλυση αυτού του ζητήματος.

45| Ταυτοποίηση των αντιδράσεων που είναι υπεύθυνες για τη σύνθεση του συστατικού της βιομάζας.

Με την εντολή *printFluxVector(model,FBAsolution.x)* δίνεται η λίστα όλων των αντιδράσεων του μοντέλου και των ροών τους.

46| Για κάθε μία από αυτές τις αντιδράσεις, ακολουθείται συγκεκριμένη διαδικασία επίλυσης.

Για κάθε μία από τις αντιδράσεις, οι οποίες συμμετέχουν στη σύνθεση του συστατικού της βιομάζας που φαίνεται να μην παράγεται από το μοντέλο και οι οποίες έχουν μηδενική ροή, ακολουθείται το παρακάτω διάγραμμα της Εικ.15.



Εικόνα 5.15 Διάγραμμα ροής για τον εντοπισμό σφαλμάτων στις αντιδράσεις δικτύου που δεν μπορούν να φέρουν ροή. Ως ‘rxn’ συμβολίζεται η αντίδραση, ως ‘conf’ συμβολίζεται το confidence score και ως ‘met’ συμβολίζεται ο μεταβολίτης.

Έλεγχος για μπλοκαρισμένες αντιδράσεις.

47| Αλλαγή συνθηκών προσομείωσης σε πλούσιο μέσο ή «άνοιγμα» όλων των αντιδράσεων ανταλλαγής (exchange reactions).

`model=changeRxnBounds(model, rxnNameList,value,boundType)`

Σε αυτό το σημείο η τιμή των αντιδράσεων ανταλλαγής δεν έχει σημασία καθώς αυτό το βήμα ελέγχει μία ποιοτική και όχι ποσοτική ιδιότητα. Συνεπώς, μπορούμε να «ανοίξουμε» τα όρια των αντιδράσεων στο $-\infty$ (π.χ. -1000) και στο $+\infty$ (π.χ. +1000).

48| Ανάλυση και αναζήτηση μπλοκαρισμένων αντιδράσεων.

Λόγω προβλήματος με την εντολή `findBlockedReaction` χρησιμοποιήθηκε ο εξής κώδικας:

```
[minFlux,maxFlux] = fluxVariability(model,0);
minMax(:,2)=maxFlux;
minMax(:,1)=minFlux;
tol = 1e-10;
blockedReactions=cell(length(model.rxns),1); %this line will solve the
problem of double to cell
[minMax(:,1),minMax(:,2)];
cnt = 1;
for i=1:length(minMax)
    if (minMax(i,2) < tol && minMax(i,2) > -tol && minMax(i,1) < tol
&& minMax(i,1) > -tol)
blockedReactions(cnt) = model.rxns(i);
cnt = cnt + 1;
    end
end
```

Αυτός ο κώδικας επιστρέφει μία λίστα με όλες τις μπλοκαρισμένες αντιδράσεις του μοντέλου ('BlockedReactions').

49| Σύνδεση αντίδρασης με το υπόλοιπο δίκτυο (προαιρετικό βήμα).

Ανάλογα με τη λειτουργία που εξυπηρετεί η αντίδραση ακολουθείται το διάγραμμα ροής στην Εικ. 15. Πραγματοποιήθηκε συνδυασμός των βημάτων 43-49 καθώς και η ανάλυση κενών.

Συμπληρωματικά με την ανάλυση κενών και την εξεύρεση μπλοκαρισμένων αντιδράσεων χρησιμοποιήθηκε και ένα άλλο υπολογιστικό εργαλείο, το Structural Analysis Tool, ένα εργαλείο που περιλαμβάνεται στο πακέτο της SBW (Systems Biology Workbench). Για αυτό το εργαλείο χρειάζεται η εισαγωγή του μοντέλου σε μορφή .xml και δίνεται μία λίστα ανεξάρτητων και

εξαρτημένων οντοτήτων. Στη ομάδα των εξαρτημένων βλέπουμε μεταβολίτες οι οποίοι είτε είναι αδιέξοδοι είτε συμμετέχουν σε μπλοκαρισμένες αντιδράσεις και δίνεται και η σύνδεσή τους με άλλους εξαρτημένους μεταβολίτες. Για παράδειγμα, στο μοντέλο του *Nannochloropsis* εντοπίστηκε πρόβλημα στην παραγωγή του εικοσαπεντανοϊκού οξέος. Καμία ένδειξη στην ανάλυση κενών δεν εντόπιζε το πρόβλημα και μέσω του εργαλείου αυτού εντοπίστηκε συσχέτιση της μη παραγωγής του με άλλους μεταβολίτες σε μονοπάτι διαφορετικό της βιοσύνθεσης πολυακόρεστων λιπαρών οξέων.

Έλεγχος ικανότητας μοντέλου να παράγει γνωστά προϊόντα έκκρισης

50| Συλλογή λίστας γνωστών προϊόντων έκκρισης του οργανισμού καθώς και πληροφορίες για το μέσο ανάπτυξης.

51| Ορισμός περιορισμών για το επιθυμητό μέσο ανάπτυξης.

Για την αλλαγή των περιορισμών χρησιμοποιείται η εντολή *model=changeRxnBounds(model, rxnNameList,value,boundType)*.

Αν το μοντέλο απαιτείται να αναπτύσσεται επιπροσθέτως της παραγωγής του υποπροϊόντος, ορίζεται το κατώτατο όριο (*boundType='l'*) της αντίδρασης βιομάζας στην αντίστοιχη τιμή.

52| Ορισμός της αντίδρασης ανταλλαγής του προϊόντος έκκρισης ως αντικειμενικής συνάρτησης.

model = addRatioReaction(model, listOfRxns, ratioCoeff)

Η αντίδραση-εις που πρέπει να οριστούν ως αντικειμενικές συναρτήσεις δίνονται στο '*rxnNameList*' και λαμβάνουν το αντίστοιχο συντελεστή '*objectiveCoeff*'.

53| Μεγιστοποίηση ('max') για την καινούρια αντικειμενική συνάρτηση (ως έκκριση αναμένεται να έχει θετική τιμή ροής)

FBAsolution = optimizeCbModel(model,'max')

Αν το προϊόν μπορεί να παραχθεί (*FBAsolution.obj>0*), προχωράμε με το επόμενο υπο-προϊόν. Αν το προϊόν δεν μπορεί να παραχθεί (*FBAsolution.obj=0*), πιθανόν λείπει το αντίστοιχο μονοπάτι ή είναι ατελές και πρέπει να πραγματοποιηθεί ανάλυση κενών.

Έλεγχος ικανότητας μοντέλου να παράγει σε συγκεκριμένη αναλογία δύο προϊόντα έκκρισης.

54| Ορισμός περιορισμών για το επιθυμητό μέσο ανάπτυξης.

Για την αλλαγή των περιορισμών χρησιμοποιείται η εντολή
`model=changeRxnBounds(model, rxnNameList,value,boundType).`

55| Επιβεβαίωση ότι και τα δύο υποπροϊόντα μπορούν να παραχθούν ανεξάρτητα.

Επανάληψη των βημάτων 50-53.

56| Προσθήκη μίας γραμμής στον πίνακα S.

Προστίθεται μία γραμμή στον πίνακα S για να ενωθούν οι αντιδράσεις έκκρισης των υποπροϊόντων:

`model = addRatioReaction(model, listOfRxns, ratioCoeff)`

Οι δύο αντιδράσεις που είναι σε συγκεκριμένη αναλογία παρατίθενται στο 'listOfRxns'. Η αναλογία τους δίνεται στο 'RatioCoeff' με καταχώρηση του αντίστοιχου συνετελεστή σε πίνακα. Για παράδειγμα, η αναλογία 1:2 δίνεται ως [12]. Αν το μοντέλο πρέπει να αναπτύσσεται και να παράγει παράλληλα το υποπροϊόν, θέτουμε το κατώτερο όριο της αντίδρασης βιομάζας στην αντίστοιχη τιμή.

`model=changeRxnBounds(model, rxnNameList,value,boundType)`

57| Ορισμός ως αντικειμενική συνάρτηση μίας εκ των δύο αντιδράσεων έκκρισης των προϊόντων.

`model=changeObjective(model,rxnNameList)`

58| Μεγιστοποίηση ('max') για την καινούρια αντικειμενική συνάρτηση (ως έκκριση αναμένεται να έχει θετική τιμή ροής).

`FBAolution = optimizeCbModel(model,'max')`

Αν το προϊόν μπορεί να παραχθεί (FBAolution.obj>0), το δεύτερο υποπροϊόν μπορεί να παραχθεί στη συγκεκριμένη αναλογία. Αν το προϊόν δεν μπορεί να παραχθεί (FBAolution.obj=0 ή το πρόβλημα είναι ανέφικτο), η αναλογία

δεν μπορεί να επιτευχθεί. Μία πιθανή αιτία για αυτό είναι ότι ο οργανισμός δεν είχε βέλτιστη ανάπτυξη στις πειραματικές συνθήκες κάτω από τις οποίες καθορίστηκε η αναλογία. Ωστόσο, αν τεθεί στο βήμα 54 χαμηλότερο όριο στον ρυθμό ανάπτυξης μπορεί να προκληθεί ασυμφωνία, λόγω ανταγωνισμού άνθρακα στα υποπροϊόντα με την αντίδραση της βιομάζας.

Υπολογισμός φαινότυπων διαγραφής ενός γονιδίου (single gene deletion phenotypes).

59| Υπολογισμός φαινότυπων διαγραφής ενός γονιδίου (single gene deletion phenotypes).

Με χρήση της εντολής

[grRatio,grRateKO,grRateWT,hasEffect,delRxn,fluxSolution] = singleRxnDeletion(model,method,geneList)

Αυτή η εντολή επιτρέπει τη χρήση διαφορετικών μεθόδων ('method') για τη βελτιστοποίηση, π.χ. FBA, MOMA, ή γραμμική MOMA. Η λίστα των γονιδίων που διαγράφεται δίνεται στο 'geneList'. Αν δεν δοθεί λίστα γονιδίων, όλα τα γονίδια του μοντέλου θα διαγραφούν και θα ελεγχθούν ως προς την ικανότητα ανάπτυξης του του μεταλλαγμένου μοντέλου. Η εντολή αυτή υπολογίζει τον ρυθμό ανάπτυξης του wild-type στελέχους ('grRateWT') για κάθε διαγραφή ('grRateKO') καθώς και τις σχετικές αναλογίες ρυθμού ανάπτυξης ('grRatio').

60| Σύγκριση με πειραματικά δεδομένα.

Η αξιολόγηση των ασυνεπειών που προκύπτουν οδηγεί σε περαιτέρω βελτίωση του μοντέλου. Επανάληψη της ανάλυσης κενών εφόσον κρίνεται αναγκαίο.

Έλεγχος ικανότητας του μοντέλου να προβλέπει το σωστό ρυθμό ανάπτυξης ή άλλες ποσοτικές ιδιότητες.

Έλεγχος ικανότητας του μοντέλου να αναπτύσσεται αρκετά γρήγορα.

61| Βελτιστοποίηση της αντίδρασης βιομάζας σε διαφορετικά μέσα ανάπτυξης και σύγκριση με πειραματικά δεδομένα. Αν το μοντέλο δεν αναπτύσσεται καθόλου, ακολουθείται η επιλογή Α. Αν το μοντέλο δεν αναπτύσσεται αρκετά γρήγορα, ακολουθείται η επιλογή Β.

Α. Το μοντέλο δεν αναπτύσσεται καθόλου.

Έλεγχος των οριακών συνθηκών. Εφόσον είναι σωστές, είναι πιθανό η κατάσταση προσομείωσης να μην υποστηρίζει την ανάπτυξη ή το δίκτυο είναι ατελές. Αν συμβαίνει το τελευταίο, επιστροφή στην ανάλυση κενών.

Β. Το μοντέλο δεν αναπτύσσεται αρκετά γρήγορα.

Έλεγχος των οριακών συνθηκών. Εφόσον είναι σωστές, είναι πολλαπλές οι πιθανότητες ύπαρξης λάθους. Συνίσταται επαλήθευση και επιβεβαίωση των περιορισμών που έχουν οριστεί στο μοντέλο. Χρήση της εντολής *printConstraints(model,MinInf,MaxInf)*, λαμβάνουμε τη λίστα των περιορισμένων αντιδράσεων που είναι μεγαλύτερες μιας ελάχιστης τιμής και μικρότερες μιας μέγιστης τιμής.

62| Έλεγχος αν κάποιος παράγοντας του μέσου είναι περιοριστικός για την ανάπτυξη. Αύξηση του ρυθμού πρόσληψης ('value') ενός υποστρώματος ('rxnNameList') κάθε φορά με χρήση της εντολής:

model=changeRxnBounds(model, rxnNameList,value,boundType) και ορίζοντας τύπο ορίου το κατώτατο, 'l' ('boundType').

63| Μεγιστοποίηση για την βιομάζα.

Αν η τιμή της αντίδρασης βιομάζας αυξάνει, σημαίνει ότι η ένωση αυτή είναι περιοριστική. Έτσι υποδεικνύεται το σημείο στο δίκτυο που κάτι πρέπει να λείπει.

64| Προσδιορισμός του μειωμένου κόστους που σχετίζεται με αντιδράσεις του δικτύου όταν γίνεται βελτιστοποίηση της αντικειμενικής συνάρτησης.

Χρησιμοποιείται η εντολή `FBAolution = optimizeCbModel(model, osenseStr, minNorm, allowLoops)`. Εύρεση εκείνων των αντιδράσεων με τις χαμηλότερες τιμές μειωμένου κόστους. Με αύξηση της ροής διαμέσου αυτών των αντιδράσεων (απελευθέρωση του άνω ορίου) θα αυξηθεί και η ροή διαμέσου της αντικειμενικής συνάρτησης.

Έλεγχος αν το μοντέλο αναπτύσσεται υπερβολικά γρήγορα.

65| Βελτιστοποίηση της αντίδρασης βιομάζας σε διαφορετικά μέσα ανάπτυξης και σύγκριση με πειραματικά δεδομένα.

66| Επιβεβαίωση των συνθηκών περιορισμού του μοντέλου.

Μέσω της εντολής `printConstraints(model, MinInf, MaxInf)`, λαμβάνουμε τη λίστα των περιορισμένων αντιδράσεων που είναι μεγαλύτερες μιας ελάχιστης τιμής και μικρότερες μιας μέγιστης τιμής.

Πραγματοποιούνται ένας ή περισσότεροι από τους παρακάτω ελέγχους, για αναγνώριση πιθανών λαθών στο δίκτυο:

67| Επαλήθευση ότι τα κλάσματα και τα συστατικά στην αντίδραση της βιομάζας είναι συνεπή ως προς τα τρέχοντα δεδομένα. Αυτό μπορεί να σχετίζεται με το ότι η GAM στην αντίδραση βιομάζας είναι λανθασμένη.

68| Ταυτοποίηση αντιδράσεων shuttling, αναζητώντας αντιδράσεις που σχετίζονται με βρόχους (loops).

69| Επανεξέταση των θερμοδυναμικών πληροφοριών που σχετίζονται με την αντίδραση, για παράδειγμα την κατεύθυνση της αντίδρασης.

70| Χρήση single reaction deletion.

Αυτή η διαδικασία επιτρέπει τον εντοπισμό των αντιδράσεων που συμβάλλουν το μοντέλο να αναπτύσσεται υπερβολικά γρήγορα. Χρησιμοποιείται η παρακάτω εντολή, θέτωντας ως μέθοδο ('method') την 'FBA' και η λίστα αντιδράσεων ('rxnList') θα πρέπει να περιέχει μία ή

περισσότερες αντιδράσεις προς διαγραφή. Εφόσον θέλουμε να ελεγχθούν όλες οι αντιδράσεις, αφήνουμε το συγκεκριμένο πεδίο κενό.

[grRatio,grRateKO,grRateWT,hasEffect,delRxn,fluxSolution] = *singleRxnDeletion(model,method,rxnList,verbFlag)*

Η εντολή αυτή επιστρέφει τον ρυθμό ανάπτυξης του wild-type στελέχους ('grRateWT') για κάθε διαγραφή ('grRateKO') καθώς και τις σχετικές αναλογίες ρυθμού ανάπτυξης ('grRatio'). Ωστόσο, είναι πολύ πιθανό πλήθος αντιδράσεων να συμβάλλουν στο πρόβλημα, οπότε και δεν εντοπίζονται από αυτή τη μέθοδο.

71| Μειωμένο κόστος.

Η ανάλυση μειωμένου κόστους μπορεί να χρησιμοποιηθεί για την ταυτοποίηση εκείνων των αντιδράσεων που μπορούν να μειώσουν το ρυθμό ανάπτυξης (θετική τιμή κόστους). Χρησιμοποιείται η εντολή:

FBA solution = optimizeCbModel(model,osenseStr, minNorm, allowLoops)

Η δομή 'FBA solution.w' επιστρέφει τα μειωμένα κόστη όλων των αντιδράσεων του δικτύου. Το μειωμένο κόστος είναι παράγωγο της αντικειμενικής συνάρτησης σε σχέση με μία εσωτερική αντίδραση με μηδενική ροή και υποδεικνύει πόσο μία συγκεκριμένη αντίδραση επηρεάζει την αντικειμενική συνάρτηση.

«Εκτύπωση» του μοντέλου από το Matlab

72| Το τελικό μοντέλο γίνεται διαθέσιμο στην επιστημονική κοινότητα σε δύο μορφές: 1. Ως φύλλο excel που περιέχει όλες τις πληροφορίες για το μοντέλο και 2. Σε μορφή SBML έτσι ώστε να είναι συμβατό και με άλλα εργαλεία-προγράμματα μοντελοποίησης. Για την εξαγωγή του μοντέλου από το Matlab σε μορφή .xls και .xml αντίστοιχα χρησιμοποιούνται οι δύο παρακάτω εντολές:

writeCbModel(model,'xls','modelName')

writeCbModel(model,'sbml','modelName')

(Thiele & Palsson, 2010)

5.3 Βιβλιογραφία

- Augustine, D., Kawaroe, M. & Sudrajat, A. O., 2014. Effect of Ethyl Methanesulfonate on Fatty Acid Characteristic of *Nannochloropsis* sp.. *International Journal of Environment and Bioenergy*, 9(3), pp. 196-204.
- Brown, M., Jeffrey, S., Volkman, J. & Dunstan, G., 1997. Nutritional properties of microalgae for mariculture. *Aquaculture*, Volume 151, pp. 315-331.
- Brown, M. R., 1991. The amino-acid and sugar composition of 16 species of microalgae used in mariculture. *Journal of Experimental Marine Biology and Ecology*, 145(1), pp. 79-99.
- Caspi, R. & SRI International, 2011. *MetaCyc Metabolic Pathway Database*. [Online] Available at: <http://metacyc.org/> [Accessed October 2014].
- Dunstan, G., Volkman, J., Barrett, S. & Garland, C., 1993. Changes in the lipid composition and maximisation of the polyunsaturated fatty acid content of three microalgae grown in mass culture. *Journal of Applied Phycology*, 5(1), pp. 71-83.
- Gomes de Oliveira Dal'Molin, C., Quek, L.-E., Palfreyman, R. W. & Nielsen, L. K., 2011. *AlgaGEM – a genome-scale metabolic reconstruction of algae based on the Chlamydomonas reinhardtii genome*. Ouro Preto, Brazil, BMC Genomics.
- Hamana, K., Sakamoto, A., Nishina, M. & Niitsu, M., 2004. Cellular polyamine profile of the phyla Dinophyta, Apicomplexa, Ciliophora, Euglenozoa, Cercozoa and Heterokonta. *The Journal of General and Applied Microbiology*, 50(5), pp. 297-303.
- James Hutton Institute, 2014. *The AOCS Lipid Library*. [Online] Available at: <http://lipidlibrary.aocs.org/> [Accessed November 2014].
- Lodish, H. et al., 2000. Processing of rRNA and tRNA. In: *Molecular Cell Biology*. New York: W. H. Freeman.
- Lubián, L. M. et al., 2000. *Nannochloropsis* (Eustigmatophyceae) as source of commercially valuable pigments. *Journal of Applied Phycology*, 12(3-5), pp. 249-255.
- Montagud, A. et al., 2010. Reconstruction and analysis of genome-scale metabolic model of a photosynthetic bacterium. *BMC Systems Biology*, Volume 4, p. Supplementary Information.
- National Center for Biotechnology Information, 2014. *GenBank*. [Online] Available at: <http://www.ncbi.nlm.nih.gov/genbank/> [Accessed November 2014].
- Nuzzo, G. et al., 2013. Composition and Quantitation of Microalgal Lipids by ERETIC 1H NMR Method. *Marine Drugs*, 11(10), p. 3742–3753.

Reboloso-Fuentes, . M. et al., 2001. Biomass Nutrient Profiles of the Microalga *Nannochloropsis*. *Journal of Agricultural and Food Chemistry*, 49(6), pp. 2966-2972.

Sablok, G. et al., 2013. *ChloromitoSSRDB 2.00*. [Online]
Available at: <http://43.250.40.48:90/chloromitosrdb/>
[Accessed October 2014].

Schneider, J. C. & Roessler, P., 1994. Radiolabeling studies of lipids and fatty acids in *Nannochloropsis* (Eustigmatophyceae), an oleaginous marine alga. *Journal of Phycology*, Volume 30, pp. 594-598.

Sukenik, A., 1999. Production of eicosapentaenoic acid by the marine eustigmatophyte *Nannochloropsis*. In: Z. Cohen, ed. *Chemicals from Microalgae*. London, UK: Taylor & Francis, pp. 41-53.

The CRIBI Biotechnology Center, 2013. *CRIBI Genomics Nannochloropsis Genome Portal, University of Padua*. [Online]
Available at: <http://www.Nannochloropsis.org/>
[Accessed December 2014].

Thiele, I. & Palsson, B. Ø., 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, 5(1), pp. 93-121.

ΚΕΦΑΛΑΙΟ 6

ΕΦΑΡΜΟΓΗ ΑΝΑΛΥΣΗΣ ΡΟΩΝ ΙΣΟΡΡΟΠΙΑΣ ΚΑΙ ΑΝΑΛΥΣΗΣ ΕΥΡΩΣΤΙΑΣ

6.1 Αποτελέσματα

Μετά την κατασκευή και αξιολόγηση, το μοντέλο χρησιμοποιήθηκε ως πλαίσιο για τον έλεγχο της βέλτιστης λειτουργίας του κυττάρου και μέγιστης παραγωγής βιοπροϊόντων κάτω από μεικτές συνθήκες ανάπτυξης, δηλαδή χρησιμοποιώντας διαφορετικές πηγές ενέργειας και άνθρακα (mixotrophic). Εφαρμόστηκαν οι απαραίτητοι περιορισμοί και πραγματοποιήθηκε ανάλυση ισορροπίας ροών και ανάλυση ευρωστίας στο MATLAB μέσω του COBRA Toolbox. Το τελικό μοντέλο αξιολογήθηκε μέσω υπολογισμού της κατανομής των ροών σε μεικτές συνθήκες ανάπτυξης. Οι ελάχιστοι περιορισμοί που πραγματοποιήθηκαν δίνονται στον Πίνακα 6.1, και αναπαριστούν τις συνθήκες ανάπτυξης.

Πίνακας 6.1 Ελάχιστοι περιορισμοί των συνθηκών ανάπτυξης

Είσοδοι, έξοδοι και περιορισμοί	Μιξοτροφικές συνθήκες ανάπτυξης(mixotrophic)
Πηγή άνθρακα: πρόσληψη CO ₂	+
Πηγή άνθρακα: πρόσληψη οξικού άλατος (acetate)	+
Πρόσληψη φωτονίων (ελεύθερη ροή)	+
Βελτιστοποίηση: μεγιστοποίηση παραγωγής	Βιομάζας
Ρυθμός βιομάζας [σε hr ⁻¹]	0,0299

Οι κατανομές των ροών υπολογίστηκαν με χρήση γραμμικού προγραμματισμού:

maximize παραγωγή βιομάζας

subject to $S \cdot v = 0$

$$V_{\text{biomass}} = b$$

$$V_{\text{min}} \leq v \leq v_{\text{max}}$$

Αυτοί οι περιορισμοί επιτρέπουν σε θερμοδυναμικά μη εφικτούς βρόχους (βρόχοι τύπου 3) να συμβαίνουν, όπου η ροή πραγματοποιείται γύρω από ένα κύκλο χωρίς καθαρή μεταβολή μεταβολιτών, και πρέπει να εντοπιστούν και να αφαιρεθούν. Αυτό συμβαίνει με την ενσωμάτωση συστημικών θερμοδυναμικών περιορισμών ανάλογων του δεύτερου νόμου του Kirchoff στο μοντέλο. Όπως η πτώση δυναμικού γύρω από ένα κλειστό βρόχο σε ένα ηλεκτρικό κύκλωμα πρέπει να είναι μηδέν, μηδενική θα πρέπει να είναι η μεταβολή ελεύθερης ενέργειας γύρω από ένα βιοχημικό βρόχο, καθώς η ελεύθερη ενέργεια είναι μεταβλητή κατάστασης. Για παράδειγμα, θα πρέπει κάθε σειρά βημάτων αντίδρασης που μαζί αποτελούν μία συνολική αντίδραση τύπου $A + 2B + C \rightarrow A + 2B + C$ θα πρέπει να έχει μεταβολή ελεύθερης ενέργειας ίση με το μηδέν. Από το δεύτερο νόμο της θερμοδυναμικής, κάθε αντίδραση στο δίκτυο θα πρέπει να έχει αρνητική μεταβολή ελεύθερης ενέργειας προς την κατεύθυνση της καθαρής ροής. (Price, et al., 2002) Η τύποι βρόχων και η εξάλειψή τους φαίνεται στην Εικ. 1.

Επομένως, για τον υπολογισμό της βέλτιστης παραγωγής βιομάζας χρησιμοποιήσαμε loopless FBA (ανάλυση ισορροπίας ροών χωρίς βρόχους) :

Μετατρέπουμε όλο το μοντέλο σε αναντίστρεπτο, έτσι ώστε όλες οι αντιδράσεις να ικανοποιούν την εξής συνθήκη:

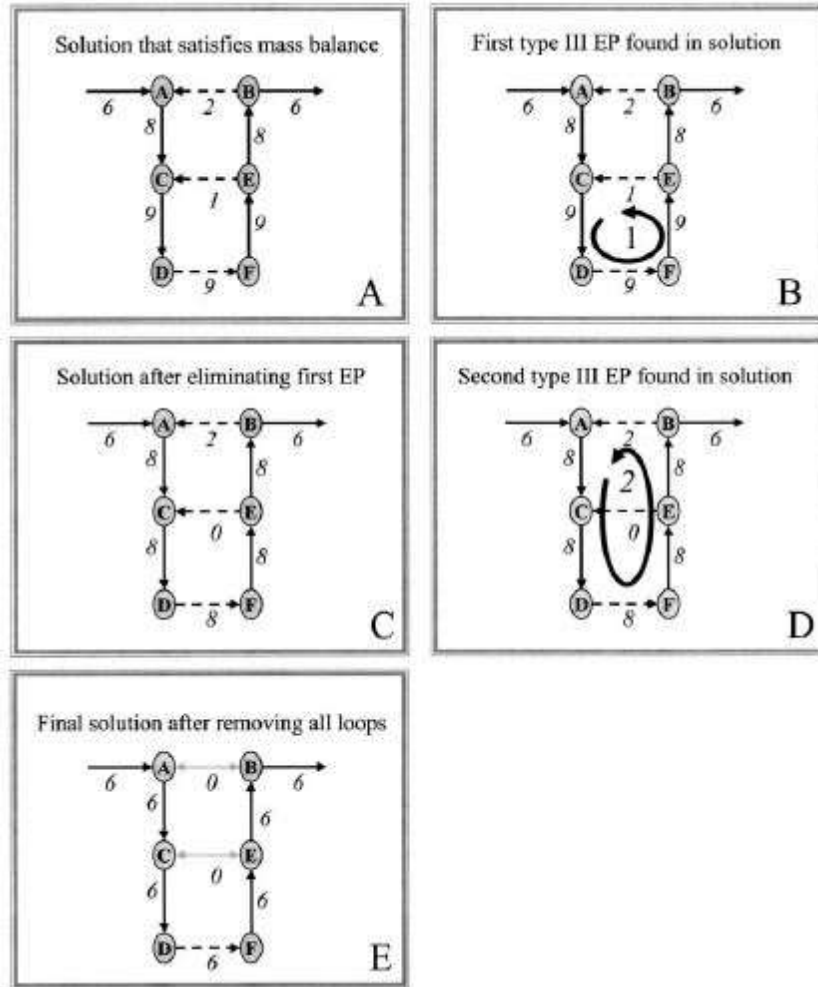
$$0 \leq lb \leq u \leq ub \leq \max(ub) ,$$

όπου lb: κάτω όριο αντίδρασης; ub: άνω όριο

Προσθέτουμε δείκτες Boolean έτσι ώστε

$$\max(ub)_i \geq v$$

$$i \in 0,1$$



Εικόνα 6.1 (Α) Μία κατανομή ροής που ικανοποιεί την ισορροπία μάζας. Ωστόσο παραβιάζεται ο κανόνας των βρόχων. (Β) Ο πρώτος τύπος ΙΙΙ ακραίου μονοπατιού [type III extreme pathways]. (C) Το αποτέλεσμα της κατανομής των ροών με εξάλειψη του κανόνα βρόχων. (D) Ένας δεύτερος τύπος ΙΙΙ ακραίου μονοπατιού. (E) Η τελική λύση όπου το ακραίο μονοπάτι (D) έχει εξαιρεθεί. Οι διακεκομμένες γραμμές υποδεικνύουν αντιστρεπτές αντιδράσεις (Price, et al., 2002)

Επιπλέον θέλουμε να εξασφαλίσουμε ότι μία εισαγωγή στο χώρο σειρών του πίνακα S υπάρχει με αρνητικές τιμές όπου v είναι μη μηδενικό. Με την έκφραση $1-i$ υποδεικνύεται η αδράνεια μιας αντίδρασης.

$$S^T x - (1 - i) (\max(ub) + 1) \leq -1$$

Τελικά:

$$\begin{pmatrix} S & 0 & 0 \\ -I & \max(ub)I & 0 \\ 0 & (\max(ub) + 1)I & S^T \end{pmatrix} \cdot \begin{pmatrix} v \\ i \\ x \end{pmatrix} = \begin{matrix} 0 \\ \geq 0 \\ \leq \max(ub) \end{matrix}$$

Η ανάλυση ισορροπίας ροών χωρίς βρόχους πραγματοποιείται μέσω της εντολής *FBA* solution = *optimizeCbModel(model,'true')*, αφού έχουμε ορίσει τη βιομάζα ως αντικειμενική συνάρτηση. (Schellenberger, et al., 2011)

Χαρακτηριστικά του μοντέλου

Το μοντέλο γονιδιακής κλίμακας του *Nannochloropsis* περιέχει 2205 αντιδράσεις, 2009 μεταβολίτες και 1695 γονίδια. Οι μεταβολίτες του κατανέμονται σε 5 κυτταρικά διαμερίσματα. Το ενεργό πεδίο του μοντέλου περιλαμβάνει τη γλυκόλυση (κυτταροπλασματική και χλωροπλαστική), τον κύκλο του κιτρικού οξέος, τη φωτοσύνθεση και οξειωτική φωσφορυλίωση, τη σύνθεση λιπαρών οξέων, τον κύκλο γλυοξυλικού οξέος, τον κύκλο φωτοαναπνοής, τη βιοσύνθεση χλωροφυλλών και καροτενοειδών, τη σύνθεση πολυαμινών και βιταμινών και αντιδράσεις ζύμωσης. Στο παρών μοντέλο δεν έχει ελεγχθεί το αν καλύπτει πλήρως το δευτερογενή μεταβολισμό, και ποια μονοπάτια αυτού είναι ενεργά.

Πενηντατρείς (53) αντιδράσεις παραγωγής συστατικών βιομάζας περιγράφουν τη συσσώρευση υδατανθράκων, σακχάρων, αμινοξέων, λιπαρών οξέων, χρωστικών, καροτενοειδών, βιταμινών και καλύπτουν σε μεγάλο βαθμό τα κυρίαρχα προϊόντα ενός κυττάρου του *Nannochloropsis*. Είκοσιεννέα (29) βασικές αντιδράσεις εναλλαγής έχουν συμπεριληφθεί για να περιγράψουν την πρόσληψη φωτός (απορροφούμενα φωτόνια), την πρόσληψη/έκκριση ανόργανων ενώσεων (CO_2 , H_2O , HCO_3^- , O_2 , NO_3 , NH_3 , H_2S , SO_4^{2-} , PO_4^{3-}), τη μετατόπιση προϊόντων ζύμωσης (όπως το οξικό άλας, το γλουταμινικό οξύ, το ασπαρτικό οξύ, την αλανίνη και τη σερίνη). Μαζί με τις αντιδράσεις των συστατικών της βιομάζας, ορίζουν το ευρύ φυσιολογικό πεδίο του μοντέλου. Προστέθηκαν ακόμη αντιδράσεις μεταφοράς μεταξύ των οργανιδίων, είτε σύμφωνα με τη βιβλιογραφία, είτε όπου δεν ήταν δυνατή εύρεση βιβλιογραφίας, πιο αυθαίρετα για λόγους συνδετικότητας του μοντέλου.

Η προσομείωση αφορά τη μεγιστοποίηση της βιομάζας σε μιξοτροφικές συνθήκες ανάπτυξης, γι' αυτό αρχικά παρουσιάζονται οι βασικές αντιδράσεις εναλλαγής και οι περιορισμοί που τους επιβλήθηκαν:

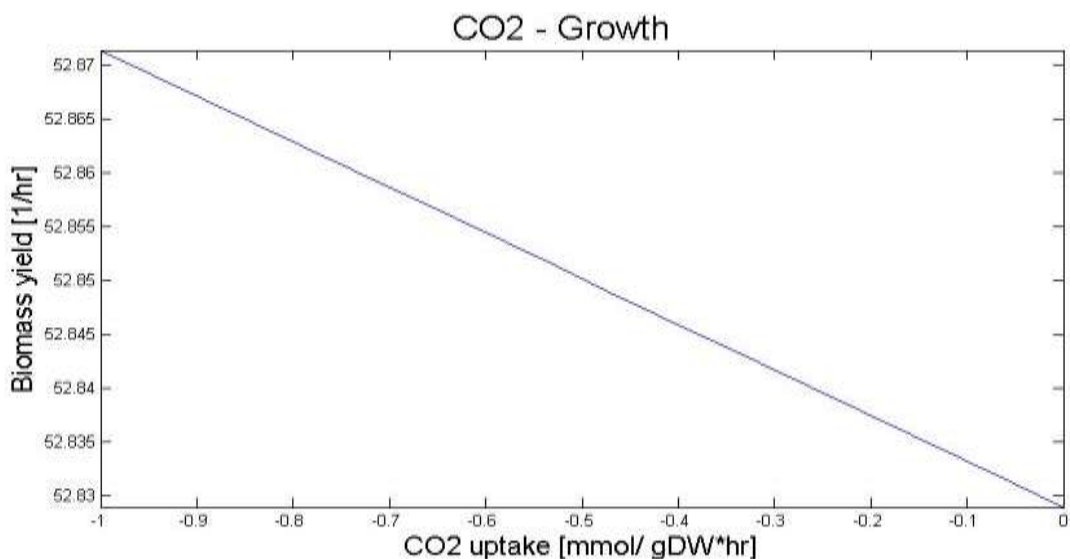
Πίνακας 6.2 Αντιδράσεις εναλλαγής και περιβαλλοντικοί περιορισμοί

	Αντίδραση	Κάτω όριο	Άνω όριο
Ex1	\Leftrightarrow S_CO2[c]	-1	-1
Ex2	\Leftrightarrow S_H2O[c]	-1000	1000
Ex3	S_Oxygen[c] \Leftrightarrow	-1000	1000
Ex4	S_Nitrate[c] \Leftrightarrow	0	0
Ex5	S_NH3[c] \Leftrightarrow	-1000	1000
Ex6	-> S_Hydrogen_32_sulfide[c]	0	1000
Ex7	\Leftrightarrow S_Sulfate[c]	-1000	1000
Ex8	\Leftrightarrow S_alpha_45_D_45_Glucose[c]	0	0
Ex10	\Leftrightarrow S_Acetate[c]	-1	-1
Ex11(φωτόνια)	-> S_hv[c]	14	28
Ex12	S_Sucrose[c] ->	0	0
Ex13	-> S_Orthophosphate[c]	0	1000
Ex14	\Leftrightarrow S_HCO3_45_[c]	0	0
Ex15	S_Glycolate[c] \Leftrightarrow	0	0
Ex16	S_Hydrogen[c] ->	0	1000
Ex17	-> S_L_45_Glutamate[c]	0	0
Ex18	-> S_L_45_Glutamine[c]	0	0
Ex19	-> S_L_45_Aspartate[c]	0	0
Ex20	-> S_L_45_Alanine[c]	0	0
Ex21	-> S_L_45_Serine[c]	0	0
Ex22	S_Formate[c] ->	0	0,001
Ex23	S_40_R_41_45_Lactate[c] ->	0	0,001
Ex24	S_Ethanol[c] ->	0	0,001
Ex27	-> S_Fe2_43_[h]	0	0,001
Ex28	-> S_Mg2_43_[h]	0	1000
Ex31	\Leftrightarrow S_Succinyl_45_CoA[c]	0	1000
Ex36	S_CO[c] ->	0	1000
Ex37	-> S_Ni2_43_[h]	-0,001	0
Ex46	S_Methanol[c] ->	0,001	0

Οι περιορισμοί για τις υπόλοιπες αντιδράσεις εναλλαγής που προστέθηκαν χάριν συνδετικότητας του μοντέλου δίνονται σε αρχείο .txt στο ηλεκτρονικό αρχείο της εργασίας. Για τις υπόλοιπες αντιδράσεις που συνθέτουν το μοντέλο τα όρια ήταν -1000 έως 1000 για αντιστρεπτές και 0 έως 1000 για αναντίστρεπτες αντιδράσεις. Επιπλέον, σε αυτό το αρχείο δίνονται και οι περιορισμοί για εκείνες τις αντιδράσεις του μοντέλου που είχαν μεγάλη ροή και αναγκαστήκαμε να τις μηδενίσουμε.

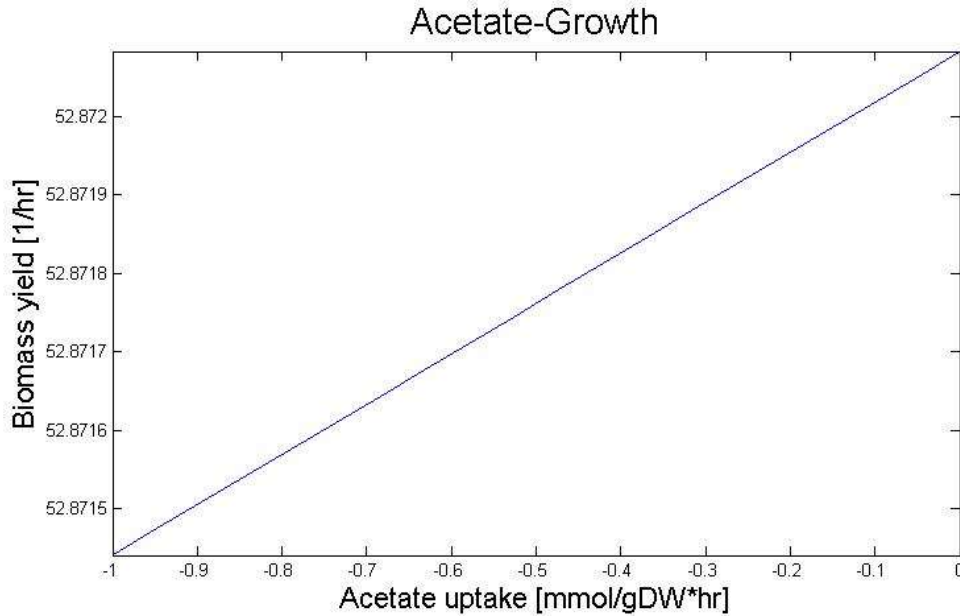
Μετά την προσομείωση αναδείχθηκε η ικανότητα του μοντέλου να αναπτύσσεται υπό ορισμένες συνθήκες, δηλαδή με διαφορετικές πηγές άνθρακα και πηγή ενέργειας το φως καθώς και η ικανότητα του να παράγει σημαντικούς μεταβολίτες.

Στο πρώτο στάδιο της προσομείωσης μελετήθηκε η ικανότητα ανάπτυξης από δύο διαφορετικές πηγές άνθρακα. Η προσομείωση πραγματοποιήθηκε με ανάλυση ευρωστίας και τα άνω όρια των αντιδράσεων εναλλαγής των υποστρωμάτων γίναν μηδέν. Στα παρακάτω διαγράμματα φαίνεται η εξάρτηση του ρυθμού ανάπτυξης από τα διάφορα υποστρώματα. Σε όλα τα διαγράμματα για μηδενική ροή υποστρώματος (της πηγής άνθρακα) παρατηρείται ανάπτυξη. Αυτό οφείλεται στο γεγονός ότι σε κάθε προσομείωση υπήρχε πάντα δεύτερη πηγή άνθρακα εκτός από αυτή που εξεταζόταν στην εκάστοτε περίπτωση. Στο Διάγραμμα 6.1 απεικονίζεται η εξάρτηση της ανάπτυξης του οργανισμού από την παροχή διοξειδίου του άνθρακα, ακολουθώντας την αναμενόμενη αύξηση της συσσώρευσης βιομάζας όσο αυξάνεται η παροχή άνθρακα. Στο Διάγραμμα 6.2 παρατηρούμε μεν μείωση της ανάπτυξης όσο



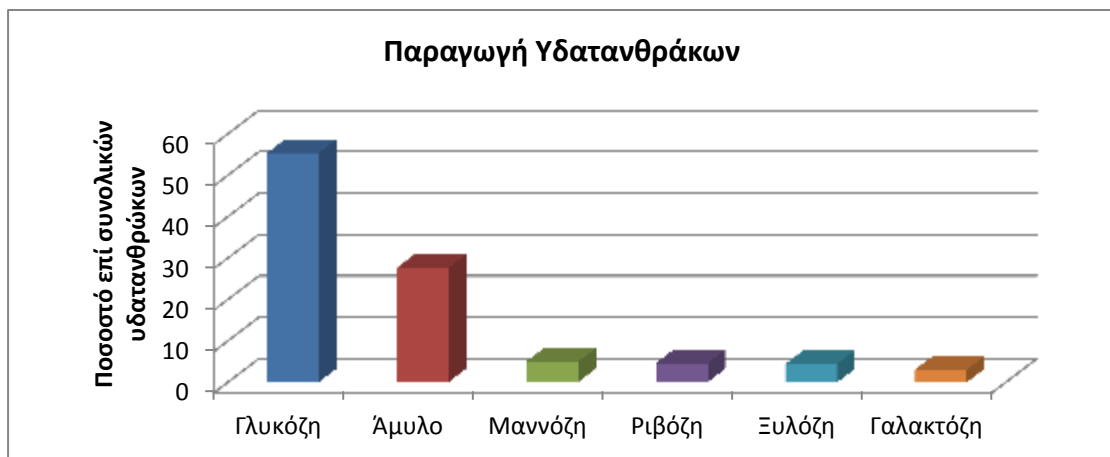
Διάγραμμα 6.1 Σχέση ρυθμού ανάπτυξης με την παροχή διοξειδίου του άνθρακα

αυξάνει η παροχή οξικού άλατος, αλλά είναι σε τόσο μικρό βαθμό που μπορεί να θεωρηθεί και σταθερή. Αυτό το αποτέλεσμα ταυτίζεται με πειραματικά αποτελέσματα τροφοδοσίας με διαφορετικές συγκεντρώσεις οξικού άλατος.



Διάγραμμα 6.2 Σχέση ρυθμού ανάπτυξης με την παροχή οξικού άλατος

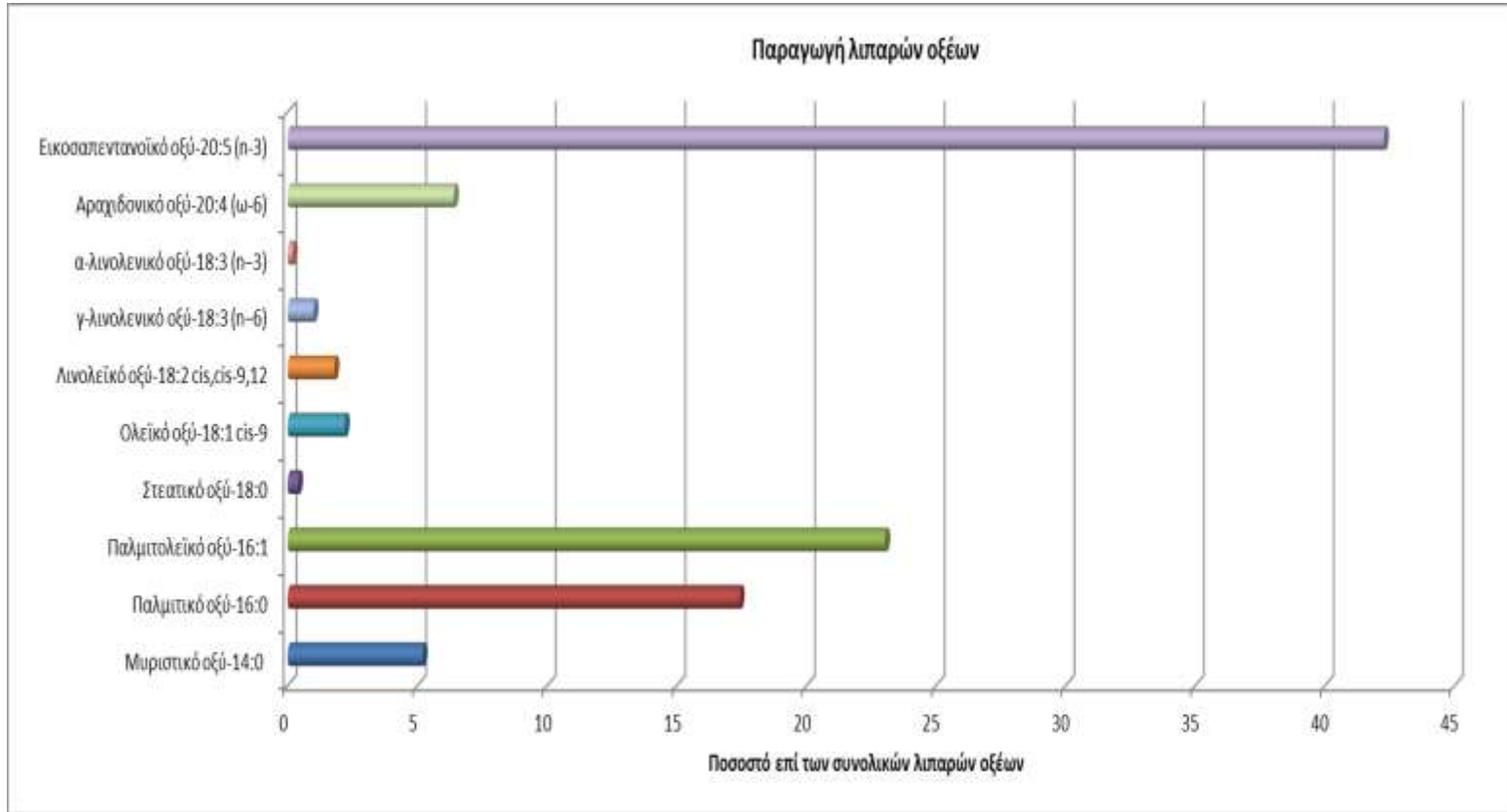
Στο δεύτερο κομμάτι της προσομείωσης ελέγχθηκε η ικανότητα παραγωγής μεταβολιτών. Η ανάλυση ισορροπίας ροών έδωσε τα εξής αποτελέσματα:



Ραβδόγραμμα 6.1 Ποσοστό παραγόμενων υδατανθράκων επι των συνολικών

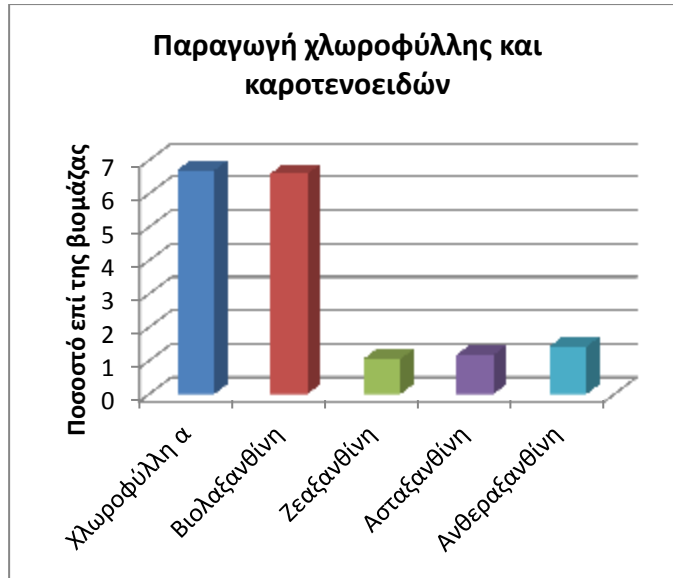
Όπως φαίνεται στο παραπάνω ραβδόγραμμα, η προσομείωση στο μοντέλο επιβεβαιώνει τη γλυκόζη ως το βασικό υδατάνθρακα, σε ποσοστό που φτάνει το 60%

επί της συνολικής συσσώρευσης υδατανθράκων. Το ποσοστό αυτό βρίσκεται πολύ κοντά αριθμητικά σε πειραματικά δεδομένα για τον οργανισμό.



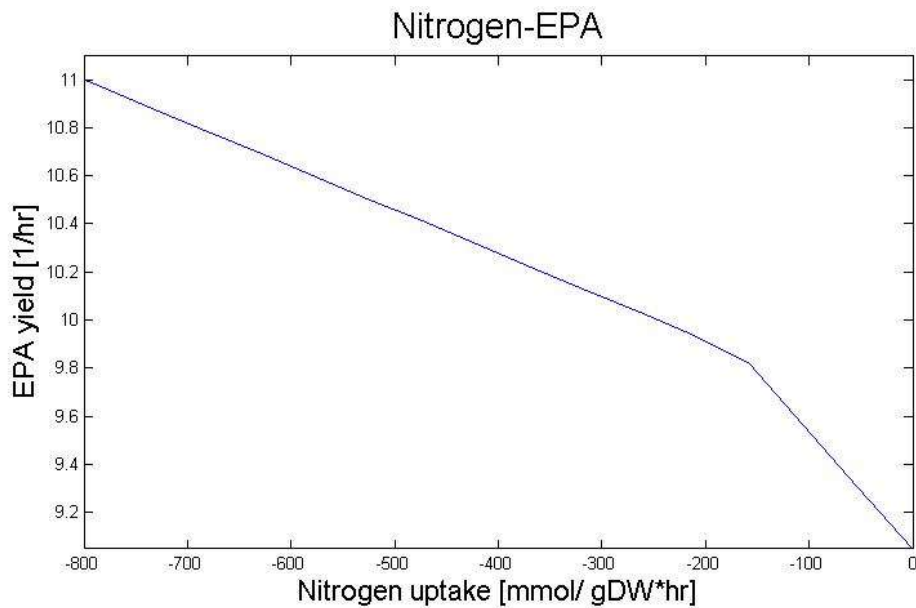
Ραβδόγραμμα 6.2 Ποσοστό παραγόμενων λιπαρών οξέων επί της συνολικής συσσώρευσης λιπαρών οξέων

Η προσομείωση αποδεικνύει ότι το μοντέλο περιλαμβάνει τα μονοπάτια σύνθεσης κορεσμένων και πολυακόρεστων λιπαρών οξέων και είναι ικανό να τα παράγει σε ποσοστά που προσομοιάζουν τις πειραματικές ενδείξεις που έχουμε. Κυρίαρχο λιπαρό οξύ είναι το εικοσαπεντανοϊκό, το οποίο και ενδιαφέρει την επιστημονική κοινότητα και τη βιομηχανία τροφίμων αλλά και γενικότερα στο πλαίσιο παραγωγής βιοντίζελ από το *Nannochloropsis*.



Ραβδόγραμμα 6.3 Ποσοτό παραγόμενων χρωστικών επί της παραγόμενης βιομάζας

Στο παραπάνω ραβδόγραμμα φαίνεται η μεγάλη παραγωγή χλωροφύλλης και βιολαξανθίνης ως ποσοστά επί της συνολικής παραγωγής βιομάζας.



Διάγραμμα 6.3 Εξάρτηση ρυθμού παραγωγής εικοσαπεντανοϊκού οξέος από το άζωτο

Στο τρίτο στάδιο της προσομείωσης ελέγχθηκε η εξάρτηση παραγωγής εικοσαπεντανοϊκού από το άζωτο. Παρατηρούμε στο Διάγραμμα 3 ότι όσο μειώνεται η παροχή αζώτου στον οργανισμό, μειώνεται η συσσώρευση του λιπαρού οξέος. Το αποτέλεσμα αυτό της προσομείωσης ποιοτικά ταυτίζεται με βιβλιογραφικά και πειραματικά δεδομένα.

6.2 Συμπεράσματα

Η κατασκευή του μοντέλου του μικροφύκου *Nannochloropsis* και η προσομείωση τόσο του ρυθμού ανάπτυξης, της παραγωγής μεταβολιτών όσο και του ρυθμού παραγωγής εικοσαπεντανοϊκού παρουσίασαν διάφορες δυσκολίες. Οι δυσκολίες αυτές οφείλονταν τόσο στη φύση των μοντέλων γονιδιακής κλίμακας και την πολυπλοκότητα τους όσο αυξάνονται σε μέγεθος, όσο και στην έλλειψη ενιαίων πειραματικών και βιοχημικών δεδομένων.

Το μοντέλο αυτό του *Nannochloropsis* δεν αποτελεί ολοκληρωμένο μοντέλο του οργανισμού και απαιτούνται συνεχείς προσπάθειες βελτίωσής του, οι οποίες θα επικεντρωθούν και στο κλείσιμο κενών κυρίως στο δευτερογενή μεταβολισμό και τα εναλλακτικά μονοπάτια ζύμωσης (τα οποία δεν είναι ακόμα επαρκώς κατανοητά σε επίπεδο οργανιδίων για τα άλγη). Το μοντέλο θα πρέπει να αναβαθμίζεται σε συνάρτηση με νέες πληροφορίες για το γονιδίωμα του οργανισμού, κυρίως όσον αφορά τη σχέση γονιδίου-πρωτεΐνης-προϊόντος. Αυτό σχετίζεται άμεσα με το γεγονός ότι το μοντέλο μοιράζεται αρκετές αντιδράσεις του AlgaGEM και περιλαμβάνει ακόμα γονίδια του *Chlamydomonas Reinhardtii* και του *Arapidopsis Thaliana*.

Ένα χαρακτηριστικό παράδειγμα αδυναμίας του μοντέλου να ανταποκριθεί μόνο στις προσομειώσεις που αφορούν την κυτταρική ανάπτυξη είναι η μη-ικανότητά του να προσομειώνει την ανάπτυξη σε λογικά πλαίσια. Αυτό φαίνεται στα διαγράμματα που παρουσιάζουν τη σχέση των υποστρωμάτων με την ανάπτυξη. Μία πρώτη παρατήρηση αφορά την αδυναμία να προσομειώνει την παραγωγή βιομάζας με μόνο μία πηγή άνθρακα. Το δεύτερο σημείο αφορά την πολύ γρήγορη ανάπτυξη που φαίνεται να υπάρχει, η οποία δεν υφίσταται στα φυσιολογικά πλαίσια αναπαραγωγής του οργανισμού. Αυτό μπορεί να οφείλεται σε πολλούς παράγοντες. Ένας από αυτούς είναι η έλλειψη ενιαίων βιοχημικών δεδομένων. Δηλαδή στις ίδιες περιβαλλοντικές συνθήκες να υπάρχουν δεδομένα για την παραγωγή βιομάζας και μεταβολιτών. Στο πλαίσιο αυτής της εργασίας, η αντίδραση της βιομάζας κατασκευάστηκε από διαφορετικά πειραματικά δεδομένα για τις διάφορες κατηγορίες προϊόντων.

Επιπλέον η γρήγορη ανάπτυξη που εμφανίζει το μοντέλο είναι πιθανό να οφείλεται σε βρόχους μέσα στο μοντέλο που το «αναγκάζουν» να αναπτύσσεται με τέτοιο ρυθμό. Είναι πολύ σημαντικό αυτοί οι βρόχοι να εντοπιστούν και να εξαλειφθούν για να βελτιωθεί η προσομειωτική του ικανότητα.

Τέλος, είναι σημαντικό να τονιστεί ότι όλες οι προσομειώσεις που πραγματοποιήθηκαν αναδεικνύουν κυρίως ποιοτικές σχέσεις. Η λειτουργία ενός μοντέλου ως προβλεπτικό εργαλείο βασίζεται στα δεδομένα εισόδου και στις έγκυρες πειραματικές μετρήσεις που βοηθούν στη σύγκριση των φαινότυπων εργαστηριακά και υπολογιστικά. Συμπερασματικά, οι ποιοτικές αυτές σχέσεις είναι μία καλή ένδειξη για τη λειτουργικότητα του μοντέλου αλλά δεν αποτελούν ποσοτικό αποτέλεσμα παρά μόνο μία βάση για περαιτέρω βελτίωση του με αξιοποίηση διάφορων πειραματικών δεδομένων.

6.3 Βιβλιογραφία

Price, N., Famili, I., Beard, D. & Palsson, B., 2002. Extreme pathways and Kirchhoff's second law. *Biophysical Journal*, 83(5), pp. 2879-82.

Schellenberger, J., Lewis, N. E. & Palsson, B. Ø., 2011. Elimination of Thermodynamically Infeasible Loops in Steady-State Metabolic Models. *Biophysical Journal*, 100(3), pp. 544-553.

ΠΑΡΑΡΤΗΜΑ

I. Πίνακες με κανόνες γραφής αντιδράσεων και μεταβολιτών

II. Χάρτες μεταβολικών μονοπατιών του μοντέλου

I. Πίνακες με κανόνες γραφής αντιδράσεων και μεταβολιτών

Πίνακας Π.Ι.1 Κανόνες γραφής αντιδράσεων

	Περιγραφή	Παράδειγμα
1.Βέλος κατεύθυνσης	Πριν το βέλος αφήνουμε 2 κενά και μετά το βέλος ένα κενό. (2 κενά) \rightarrow / \Leftarrow (1 κενό)	A \rightarrow B
2.Άθροιση μεταβολιτών	Μεταβολίτης A(1 κενό) + (1 κενό) Μεταβολίτης B	A + B \Leftarrow Γ
3.Στοιχειομετρικοί συντελεστές	Κάθε στοιχειομετρικός συντελεστής πρέπει να απέχει διάστημα ενός κενού από το μεταβολίτη και αν υπάρχει άθροιση ένα κενό από το '+' ή το βέλος	A + 2 B \rightarrow 3 Γ
4.Όνομα αντίδρασης-κυτταρικό διαμέρισμα	Το όνομα κάθε αντίδρασης (από KEGG) ακολουθείται από μία κάτω παύλα και το γράμμα-σύμβολο του κυτταρικού διαμερίσματος στο οποίο βρίσκεται	R0001_ <i>c</i> , όπου <i>c</i> : η αντίδραση βρίσκεται στο κυτταρόπλασμα <i>_h</i> : στο χλωροπλάστη <i>_m</i> : στο μιτοχόνδριο <i>_x</i> : στο υπεροξειδίσωμα <i>_p</i> : στο περίπλασμα
5.Αντιδράσεις ζήτησης	Συμβολίζονται ως 'Bc' και ένα νούμερο	Bc1 η αντίδραση ζήτησης για τη γλυκόζη
6.Αντιδράσεις ανταλλαγής	Συμβολίζονται ως 'Ex' και ένα νούμερο	Ex16 η αντίδραση ανταλλαγής για το υδρογόνο
7.Αντιδράσεις μεταφοράς	Συμβολίζονται με 3 γράμματα στη σειρά και μετά ένα νούμερο. Το πρώτο γράμμα είναι το T (μεταφορά -transport) ενώ τα άλλα δύο γράμματα δηλώνουν τα κυτταρικά διαμερίσματα μεταξύ των οποίων γίνεται η μεταφορά	TCH1 η αντίδραση μεταφοράς της γλυκόζης από το κυτταρόπλασμα στο χλωροπλάστη και αντίστροφα

Πίνακας 2 Κανόνες γραφής συντομογραφιών (abbreviations) των μεταβολιτών.

	Περιγραφή	Παράδειγμα
1.Αρχή ονομασίας μεταβολίτη	Κάθε μεταβολίτης έχει στην έναρξη του ονόματός του το ‘S_’	S_H2O για το νερό που η ονομασία του στην KeGG είναι H2O
2.Φορτία	Τα φορτία ‘+’ και ‘-’ συμβολίζονται ως ‘_43_’ και ‘_45’ αντίστοιχα	Το H ⁺ γίνεται S_H_43_ Το HCO ₃ ⁻ γίνεται S_HCO3_45
3.Πάυλα	Η παύλα μέσα στο όνομα ενός μεταβολίτη συμβολίζεται ως ‘_45_’	Το 2-Aminomuconate γίνεται S_2_45_Aminomuconate
4.Παρενθέσεις	Το άνοιγμα ‘(’ συμβολίζεται ως ‘_40_’ ενώ το κλείσιμο ‘)’ ως ‘_41_’	Το (S)-Malate γίνεται S__40_S_41__45_Malate
5.Κενό	Το κενό συμβολίζεται ως ‘_32_’	Το Vitamin B12 γίνεται S_Vitamin_32_B12
6.Κόμμα	Το κόμμα συμβολίζεται ως ‘_44_’	Το L-2,3-Dihydrodipicolinate γίνεται S_L_45_2_44_3_45_Dihydrodipicolinate
7.Αγκύλες(1)	Το άνοιγμα ‘[’ συμβολίζεται ως ‘_91_’ ενώ το κλείσιμο ως ‘_93_’	Το Malonyl-[acp] methyl ester γίνεται S_Malonyl_45__91_acp_93__32_methyl_32_ester
8.Απόστροφος	Η απόστροφος συμβολίζεται ως ‘_39_’	Το 3',5'-Cyclic AMP γίνεται S_3_39__44_5_39__45_Cyclic_32_AMP
9.Αγκύλες(2)	Το άνοιγμα ‘{’ συμβολίζεται ως ‘_123_’ ενώ το κλείσιμο ‘}’ ως ‘_125_’	
10. Το σύμβολο ->	Το σύμβολο ‘->’ συμβολίζεται ως ‘_62_’	

11. Ερωτηματικό		Το ‘;’ συμβολίζεται ως ‘_59_’
12. Άνω κάτω τελεία		Το ‘:’ συμβολίζεται ως ‘_58_’
13. Τέλος μεταβολίτη	ονόματος	<p>Στο τέλος του ονόματος κάθε μεταβολίτη τοποθετείται ‘[β]’, όπου β το σύμβολο του κυτταρικού διαμερίσματος</p> <p>Το H₂O στο κυτταρόπλασμα γίνεται S_H₂O[c], αυτή είναι ολόκληρη η ονομασία του μέσα στις αντιδράσεις.</p> <p>[h] όταν είναι στο χλωροπλάστη</p> <p>[m] όταν είναι στο μιτοχόνδριο</p> <p>[x] όταν είναι στο υπεροξειδίου σώμα</p> <p>[p] όταν είναι στο περίπλασμα</p>

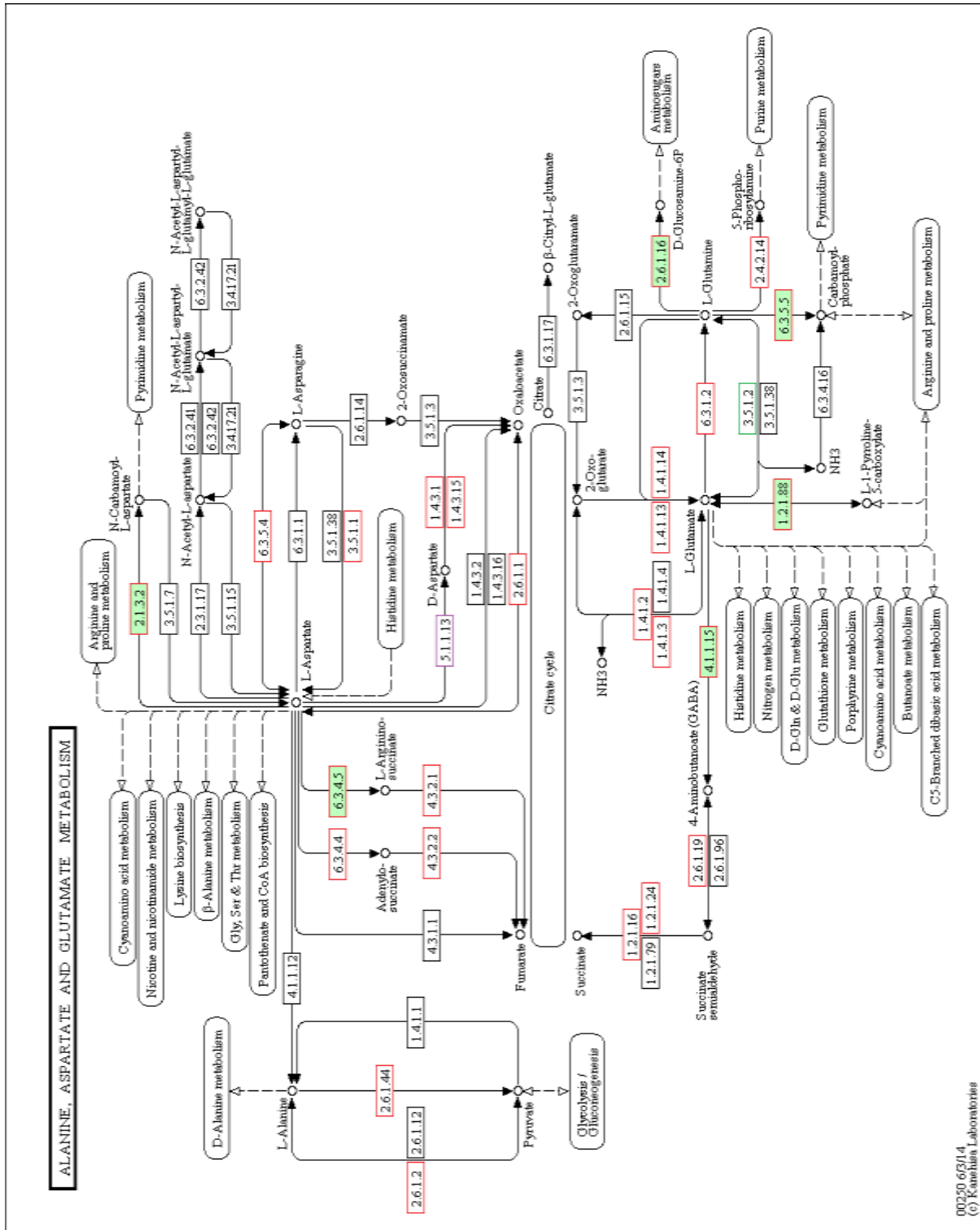
II. Χάρτες μεταβολικών μονοπατιών του μοντέλου

Σε αυτό το Παράρτημα παρουσιάζονται μερικοί μεταβολικοί χάρτες που κατασκευάστηκαν για το μοντέλο του *Nannochloropsis*, μέσω της βάσης KEGG.

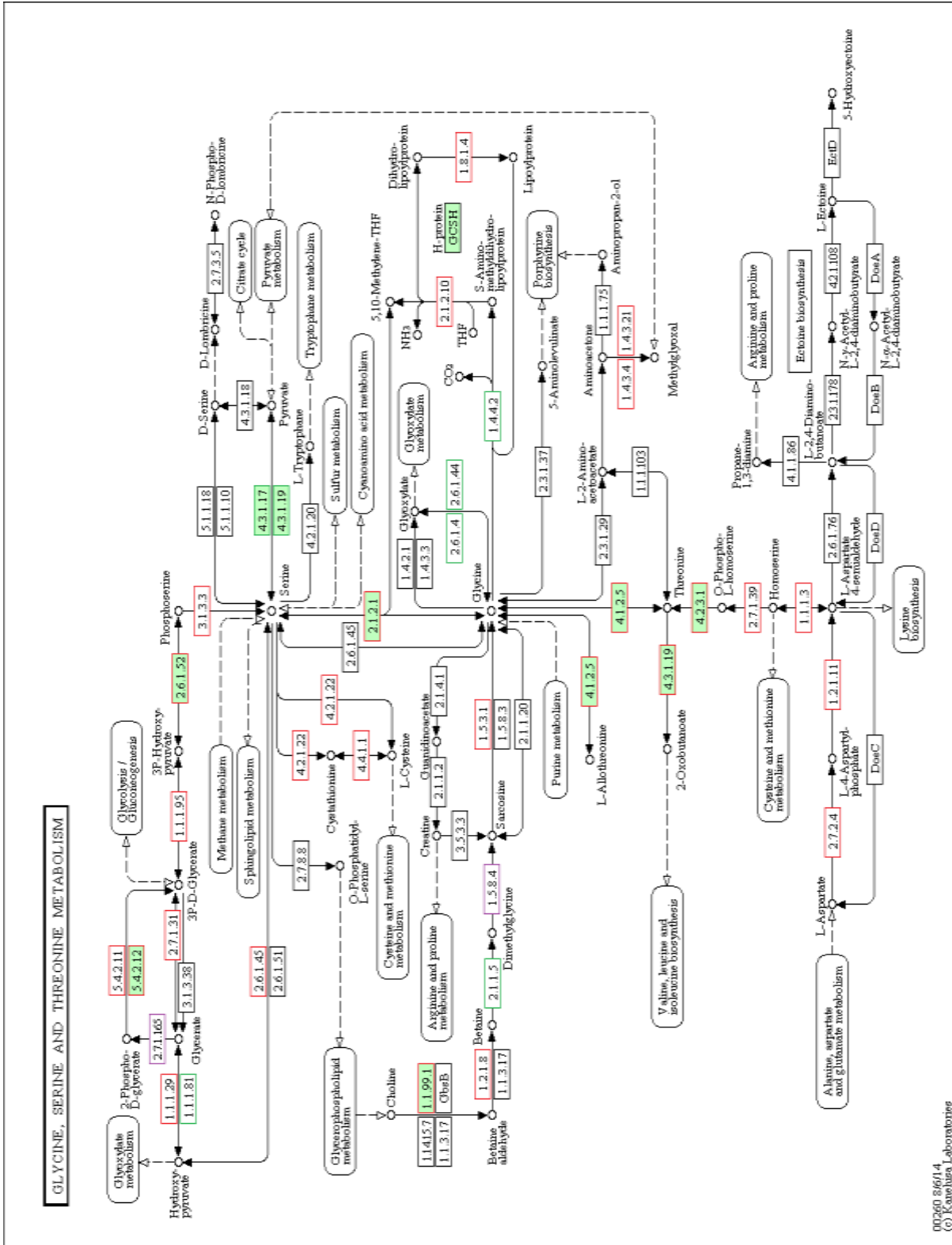
Οι χάρτες αυτοί απεικονίζουν μονοπάτια από το μεταβολισμό των αμινοξέων, των υδατανθράκων, της βιοσύνθεσης λιπιδίων, βιταμινών, τερπενοειδών και πολυκετιδίων καθώς και από το μεταβολισμό νουκλεοτιδίων και τον ενεργειακό μεταβολισμό. Οι υπόλοιποι χάρτες μαζί με αυτούς δίνονται σε ηλεκτρονική μορφή καθώς ο αριθμός τους ήταν πολύ μεγάλος για να συμπεριληφθεί στο παράρτημα.

Για την ανάγνωση των χαρτών:

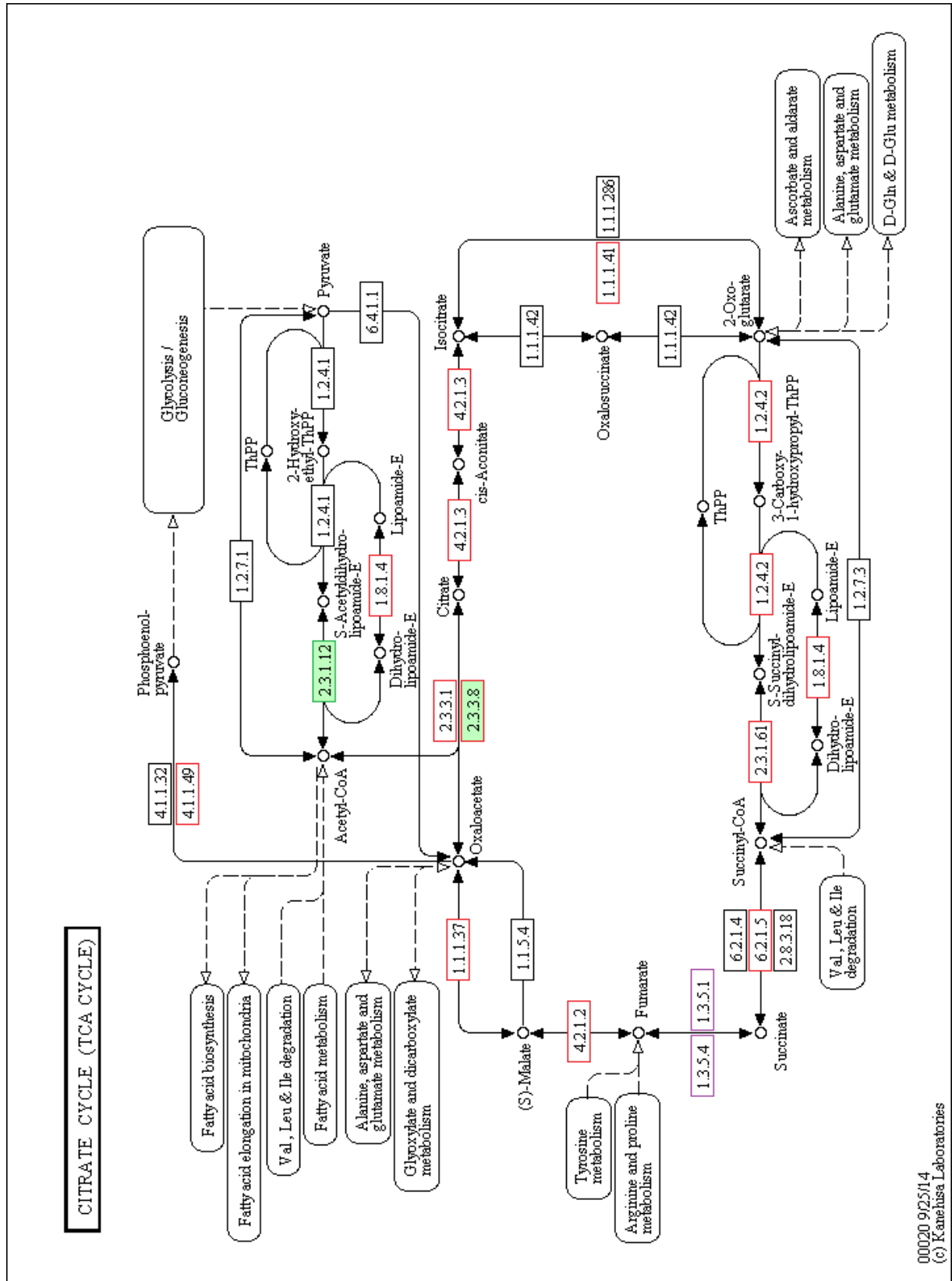
- ✓ Τα κουτιά με πράσινο χρώμα αντιπροσωπεύουν ένζυμα που είναι εξειδικευμένα για το *Nannochloropsis*.
- ✓ Τα κουτιά που έχουν κόκκινο περίγραμμα προέρχονται από το υπάρχων μοντέλο AlgaGEM
- ✓ Με πράσινο περίγραμμα εμφανίζονται τα ένζυμα για τα οποία έχει βρεθεί γονίδιο του *Nannochloropsis*
- ✓ Με μωβ περίγραμμα εμφανίζονται τα ένζυμα για τα οποία δεν βρέθηκε γονίδιο του *Nannochloropsis* (orphan reactions) και συμπληρώθηκαν κατά το στάδιο συμπλήρωσης κενών (gap filling).



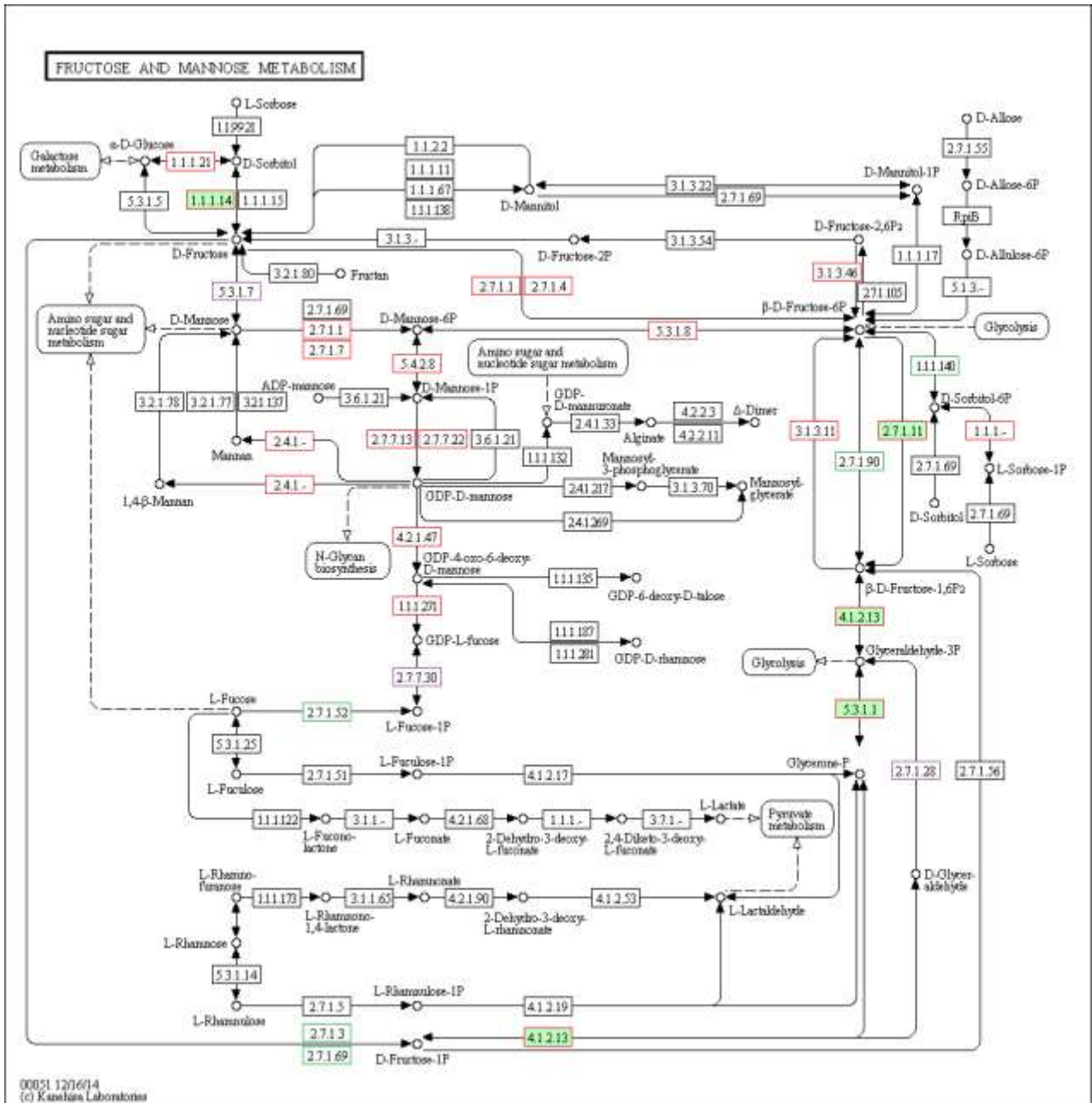
Εικόνα Π.Π.2 Μεταβολισμός αλανίνης, ασπαρτικού οξέος και γλουταμινικού οξέος



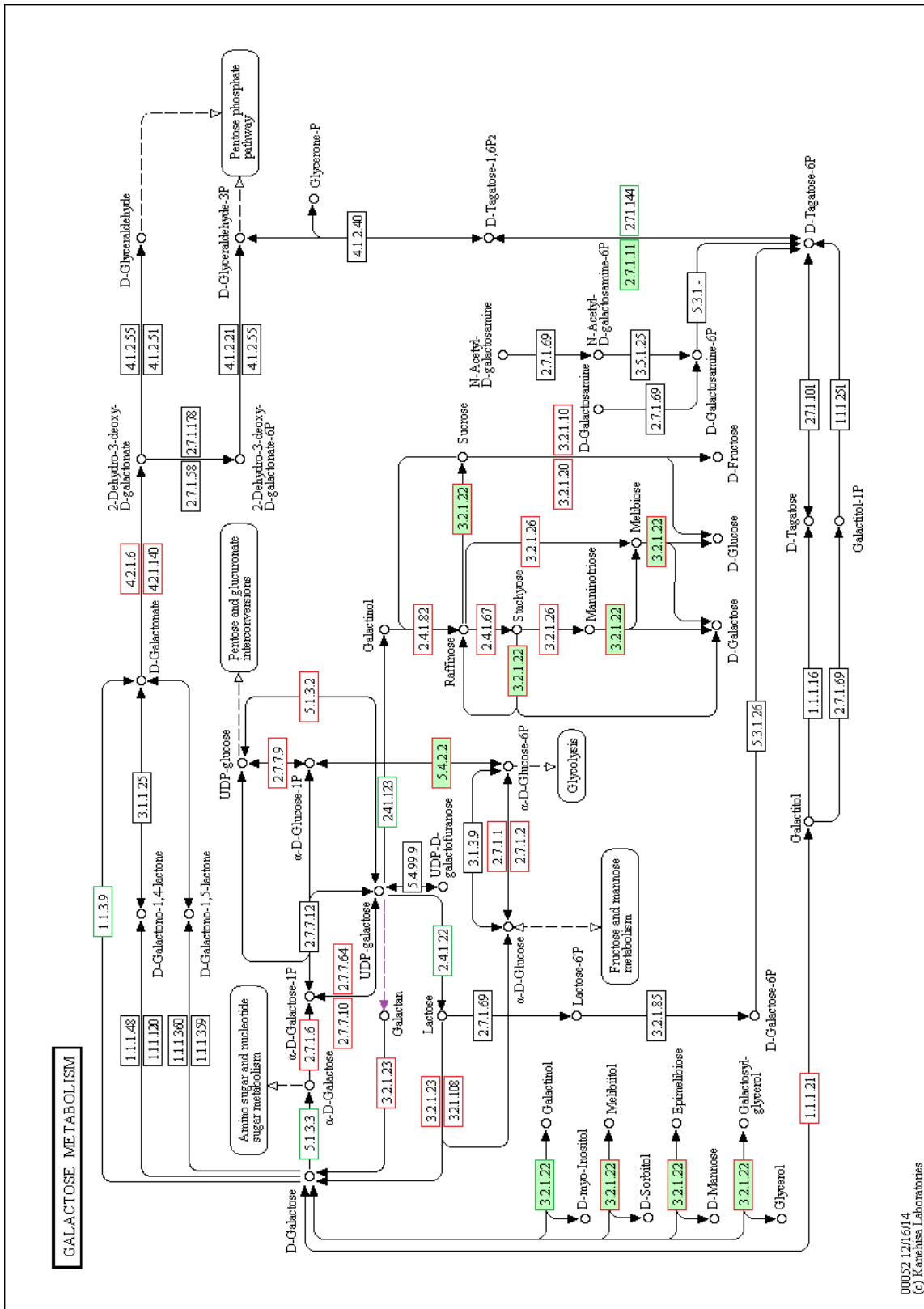
Εικόνα Π.Π.3 Μεταβολισμός γλυκίνης, σερίνης και θρεονίνης



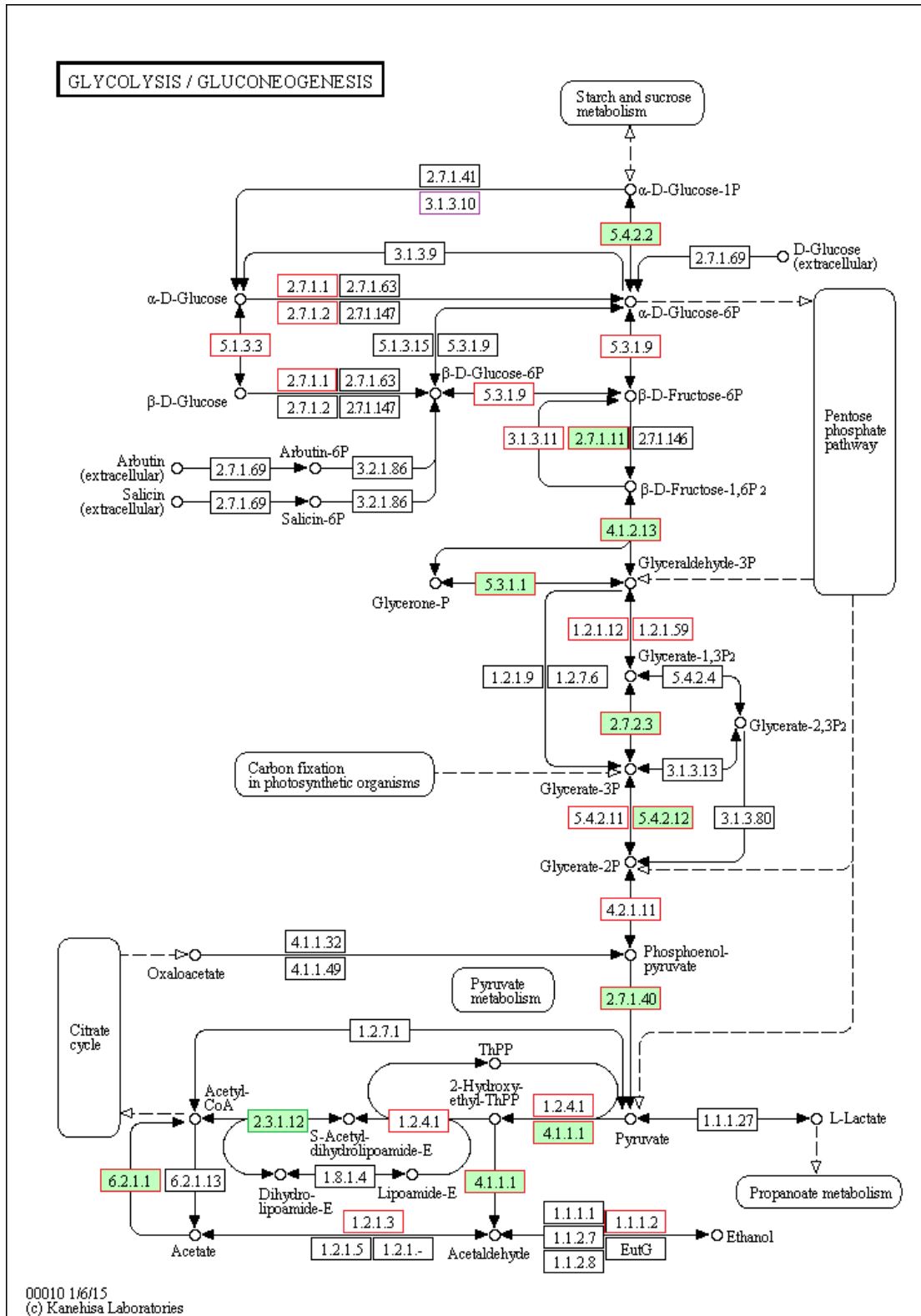
Εικόνα Π.11.4 Κύκλος Κιτρικού οξέος



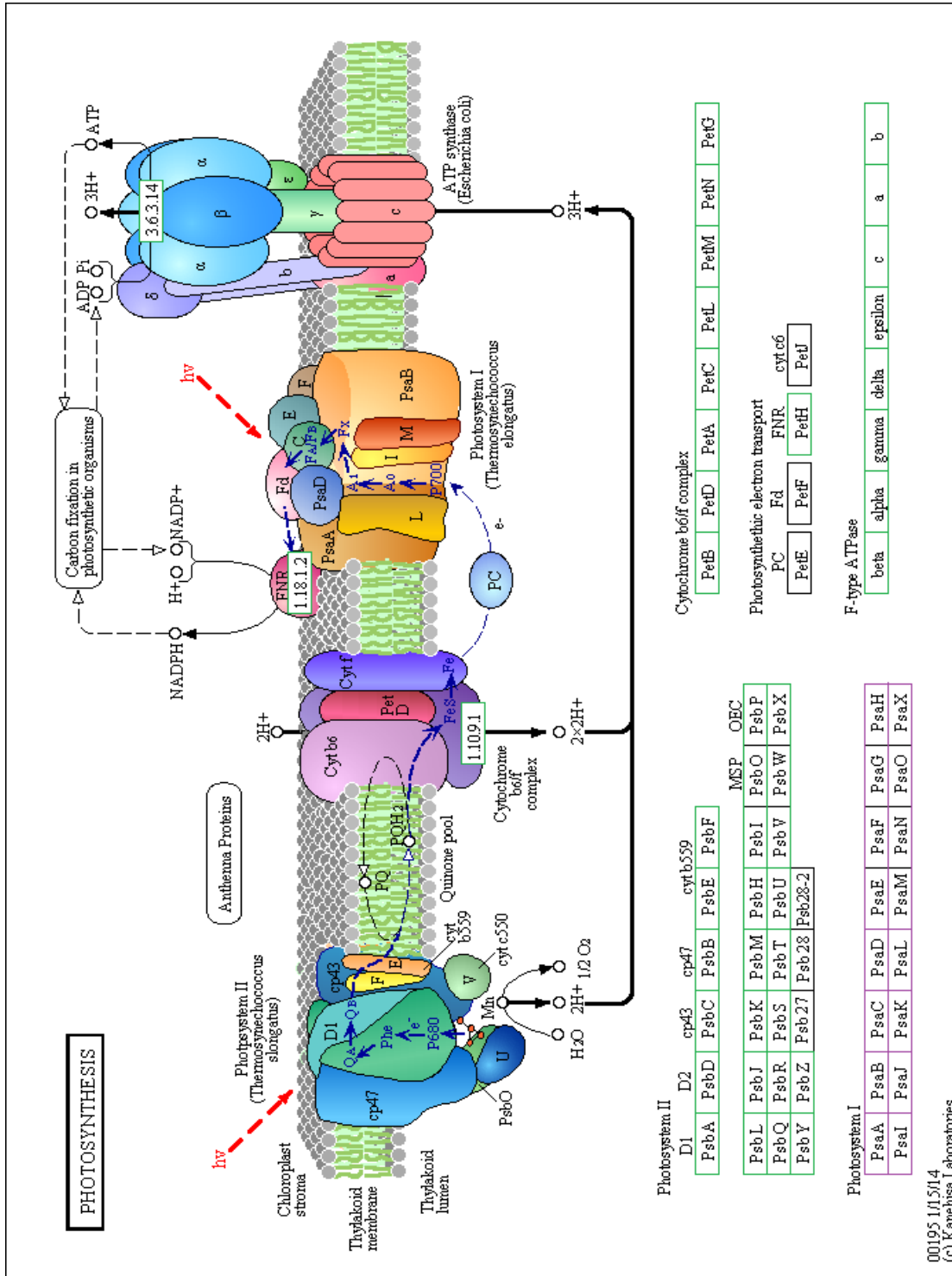
Εικόνα Π.Π.6 Μεταβολισμός γλυκόζης και μαννόζης



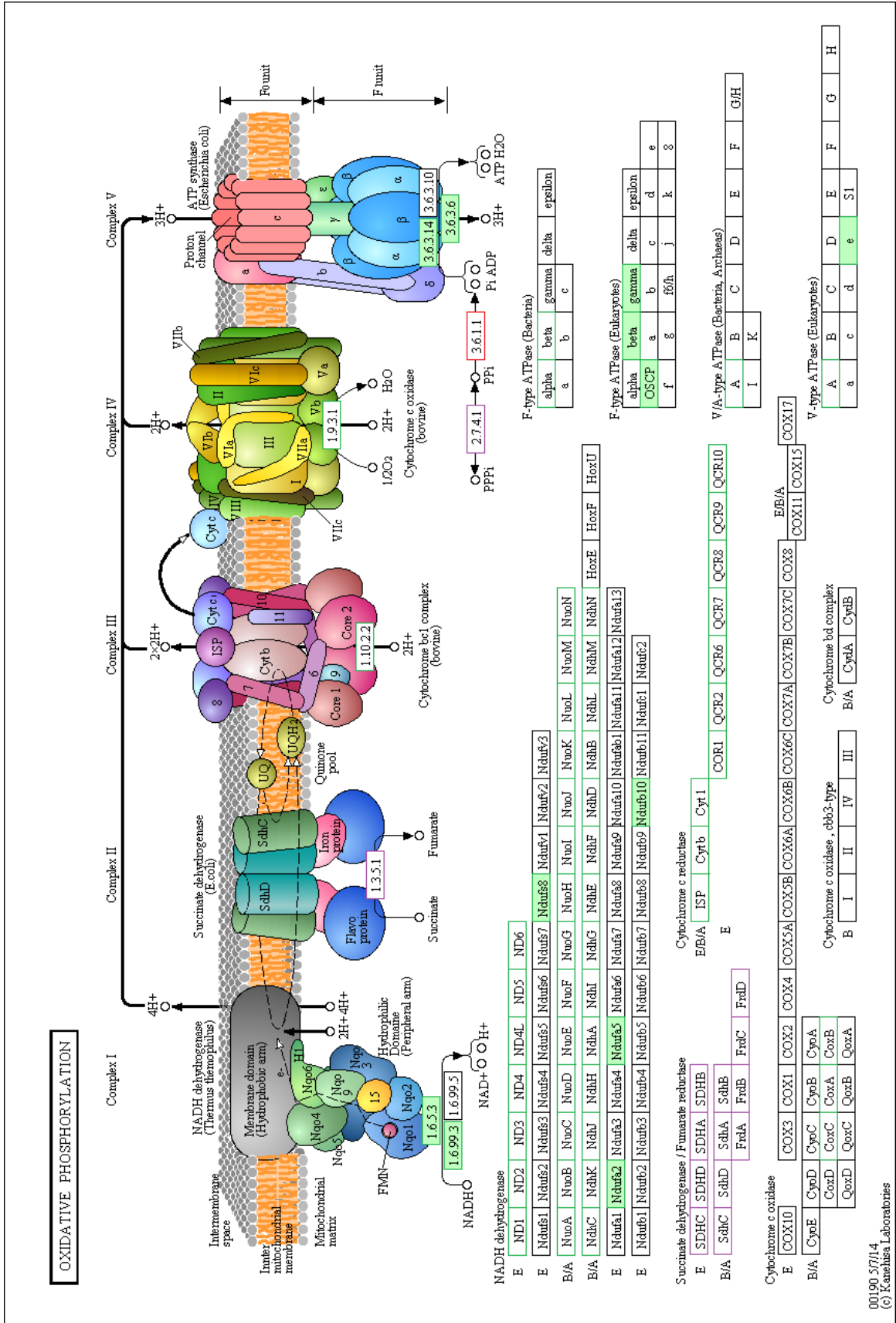
Εικόνα Π.Π.7 Μεταβολισμός γαλακτόζης



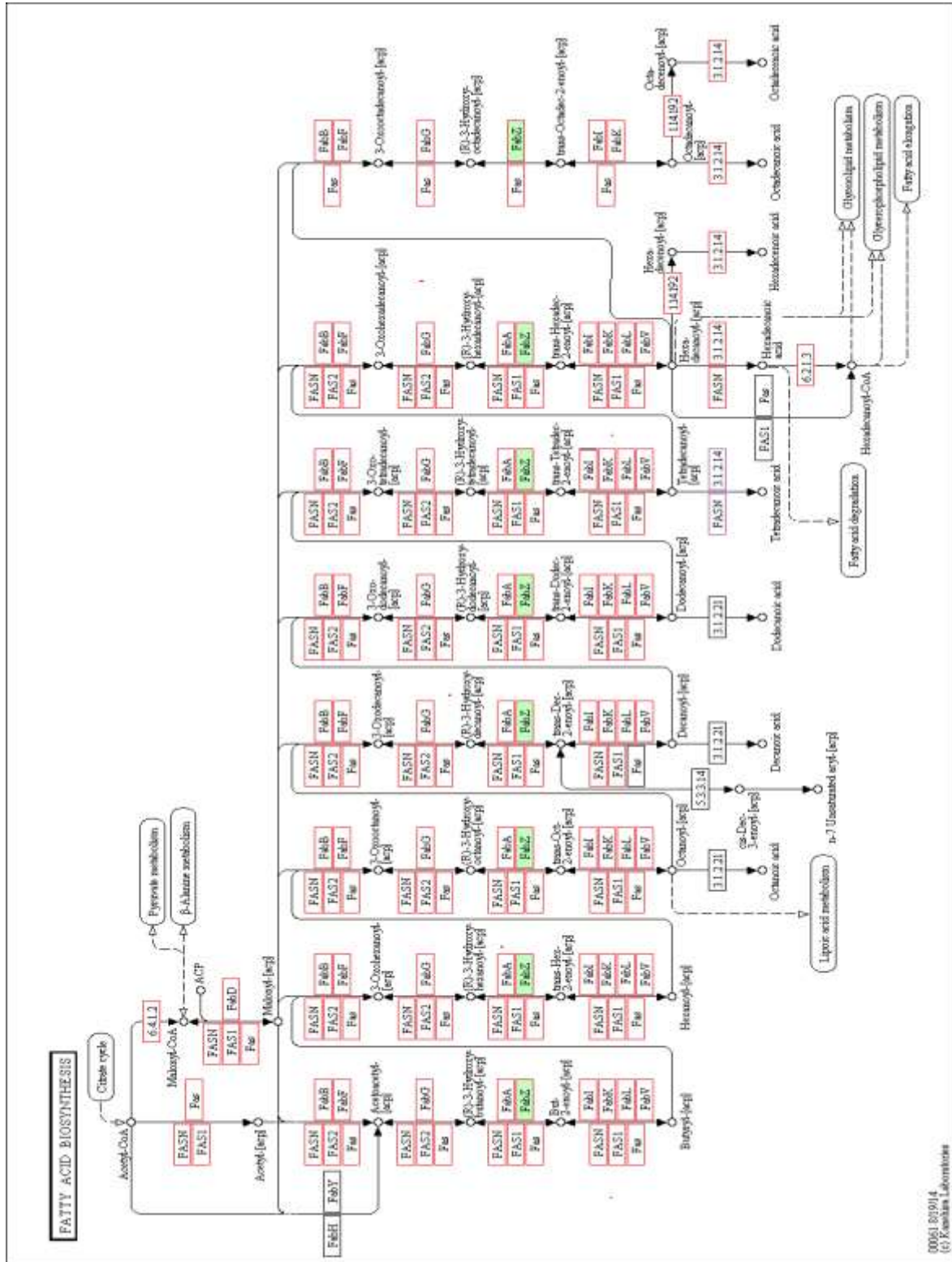
Εικόνα Π.Π.9 Μονοπάτι γλυκόλυσης



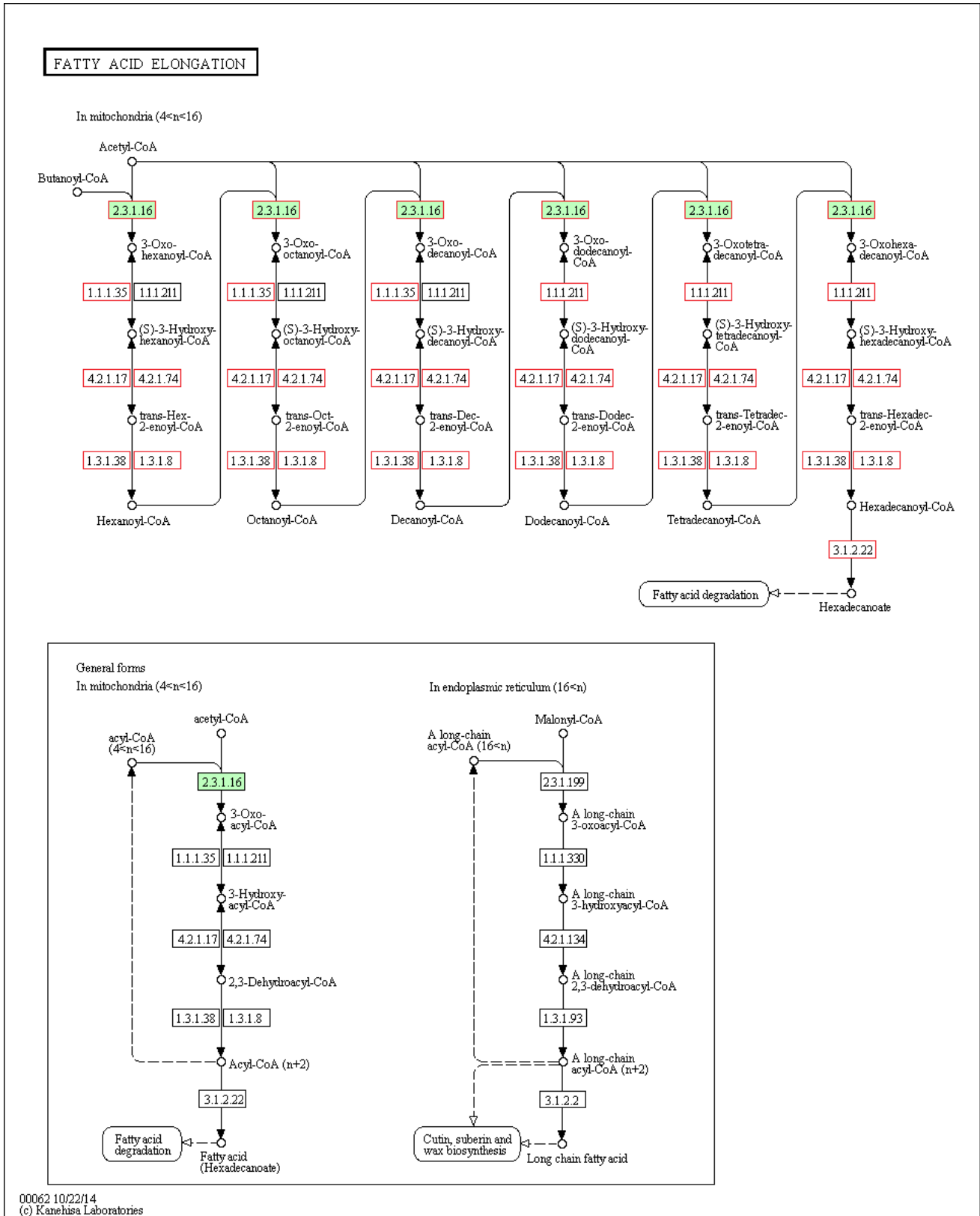
Εικόνα Π.ΙΙ.10 Φωτοσύνθεση



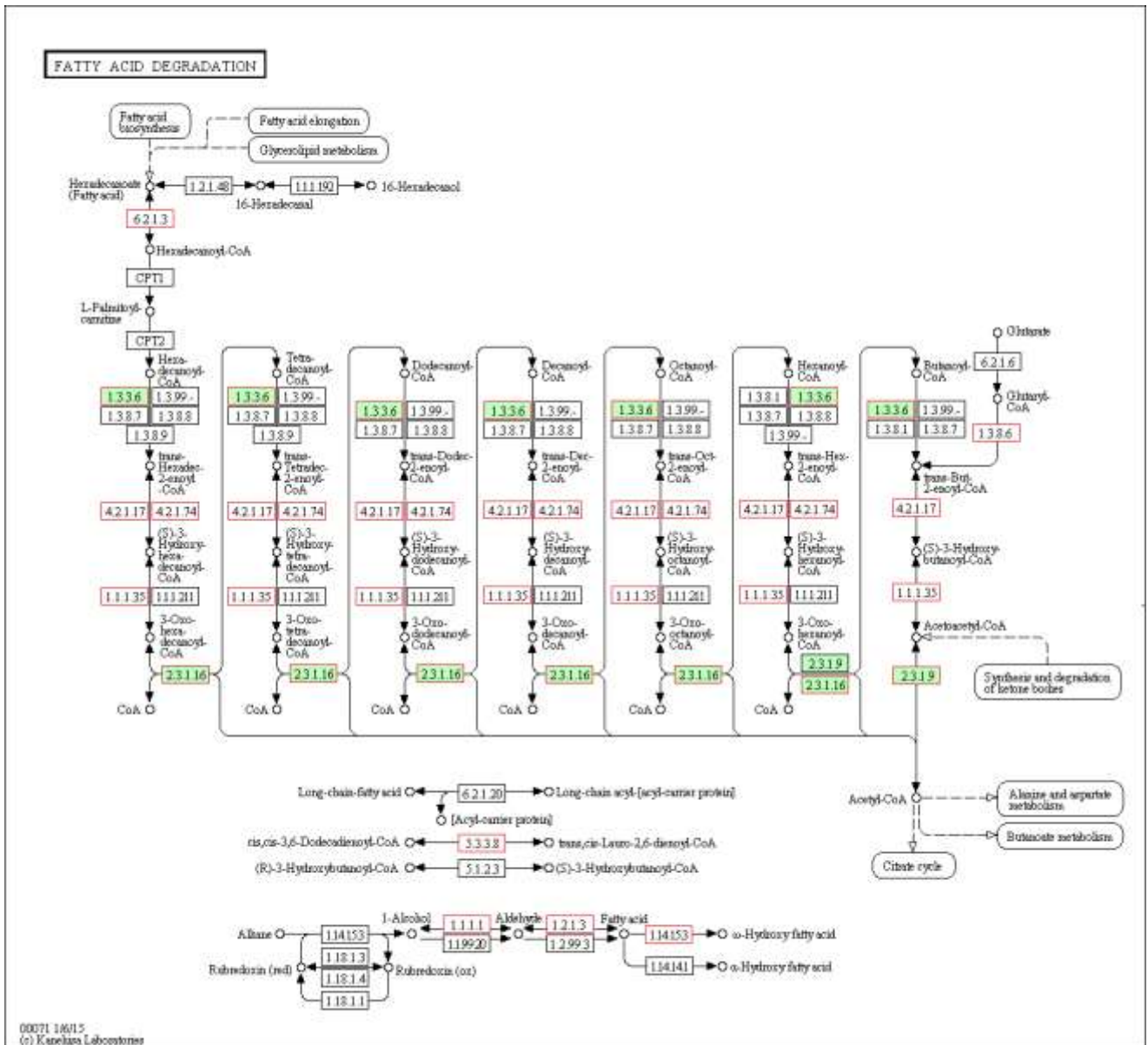
Εικόνα Π.Π.11 Οξειδωτική φωσφορυλίωση



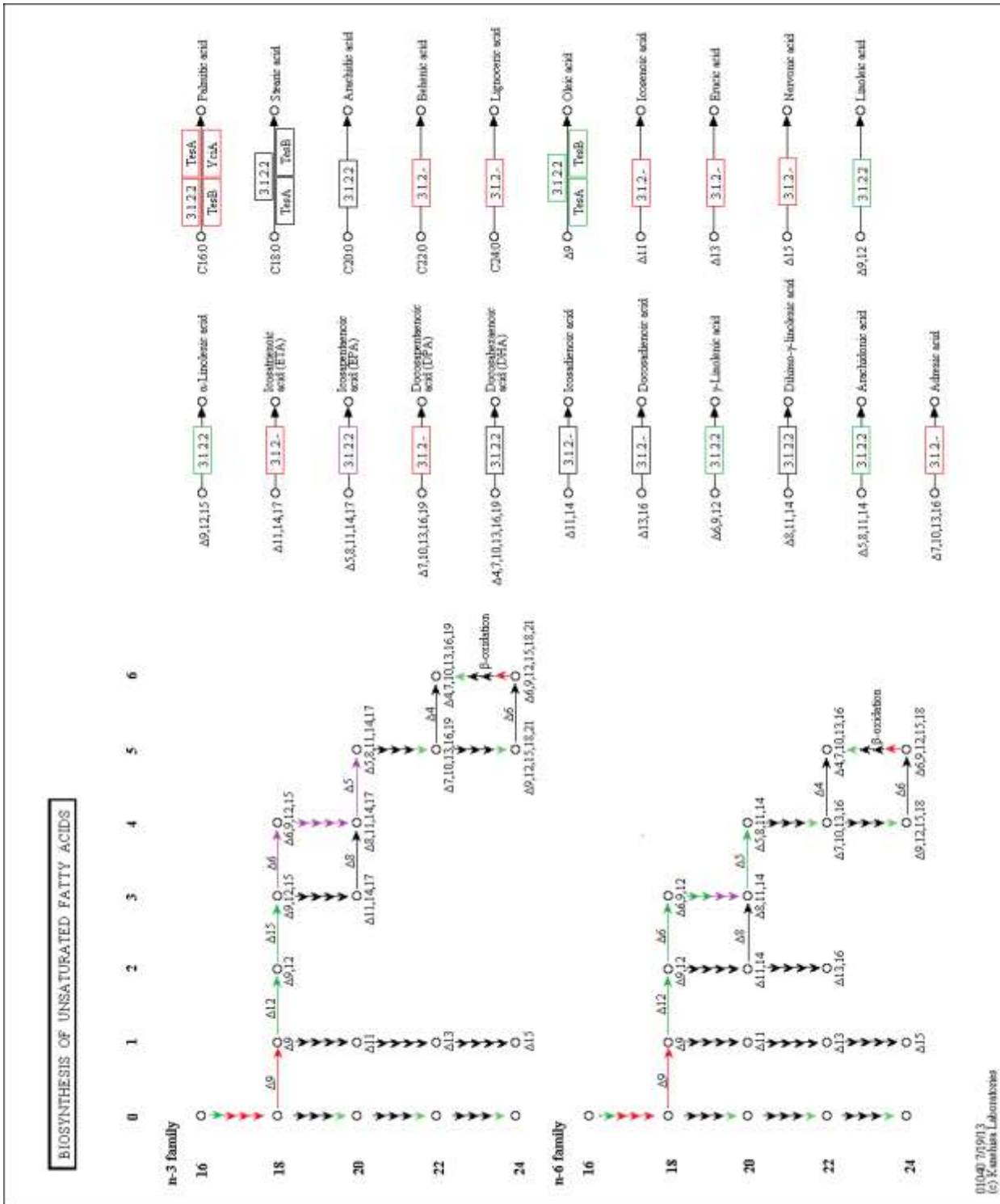
Εικόνα Π.Π.12 Βιοσύνθεση λιπαρών οξέων



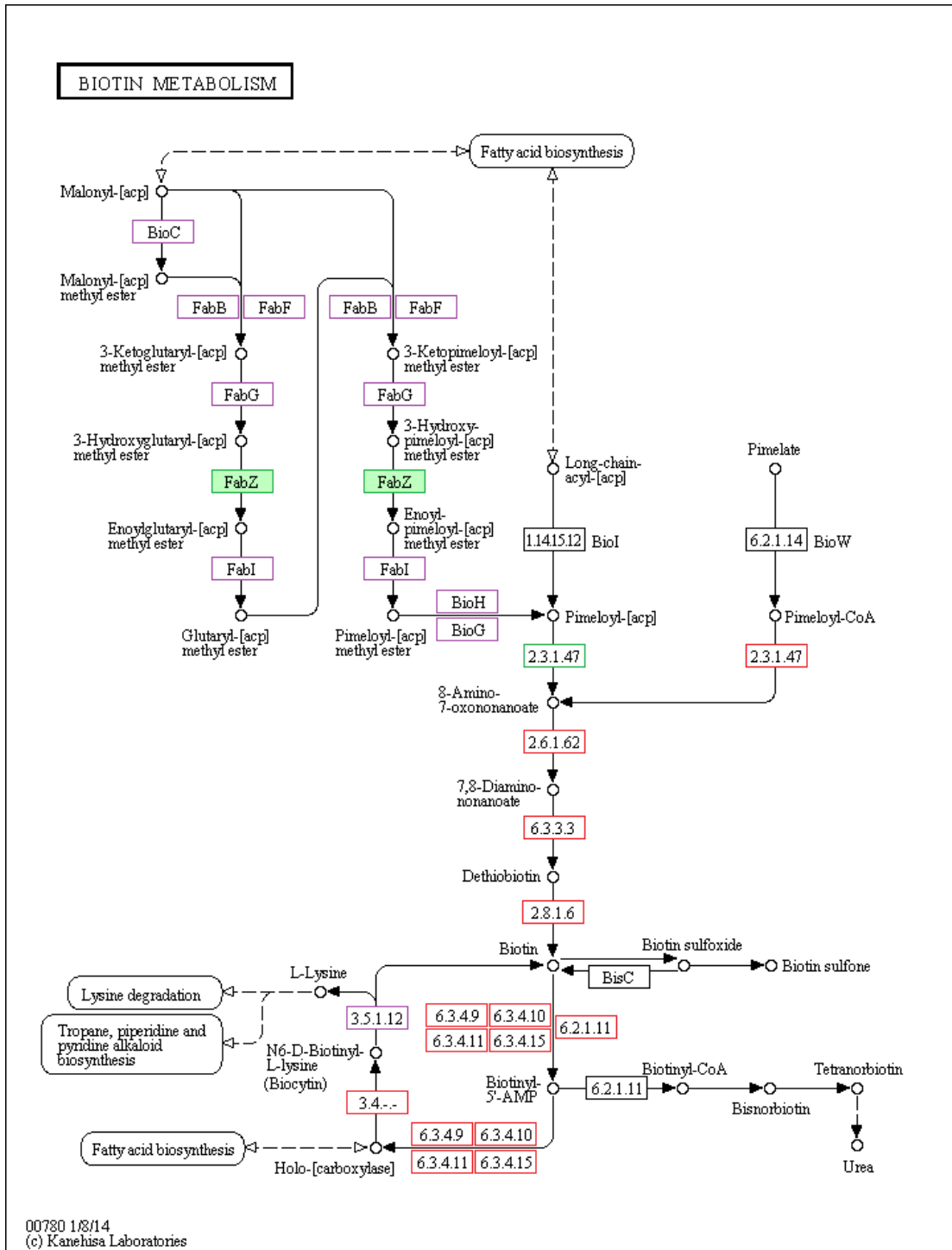
Εικόνα Π.Π.13 Επιμήκυνση λιπαρών οξέων



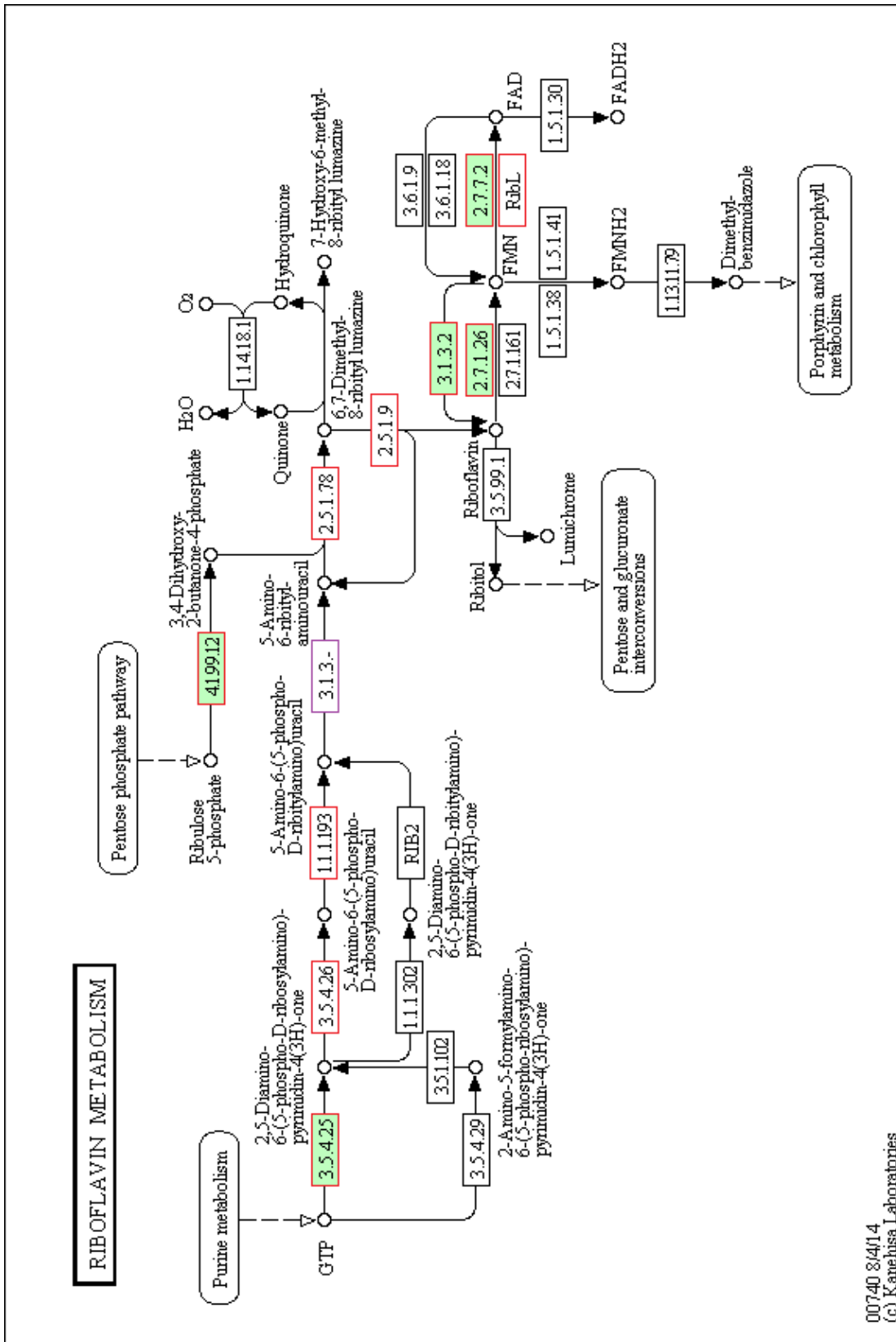
Εικόνα Π.Π.14 Αποσύνθεση λιπαρών οξέων



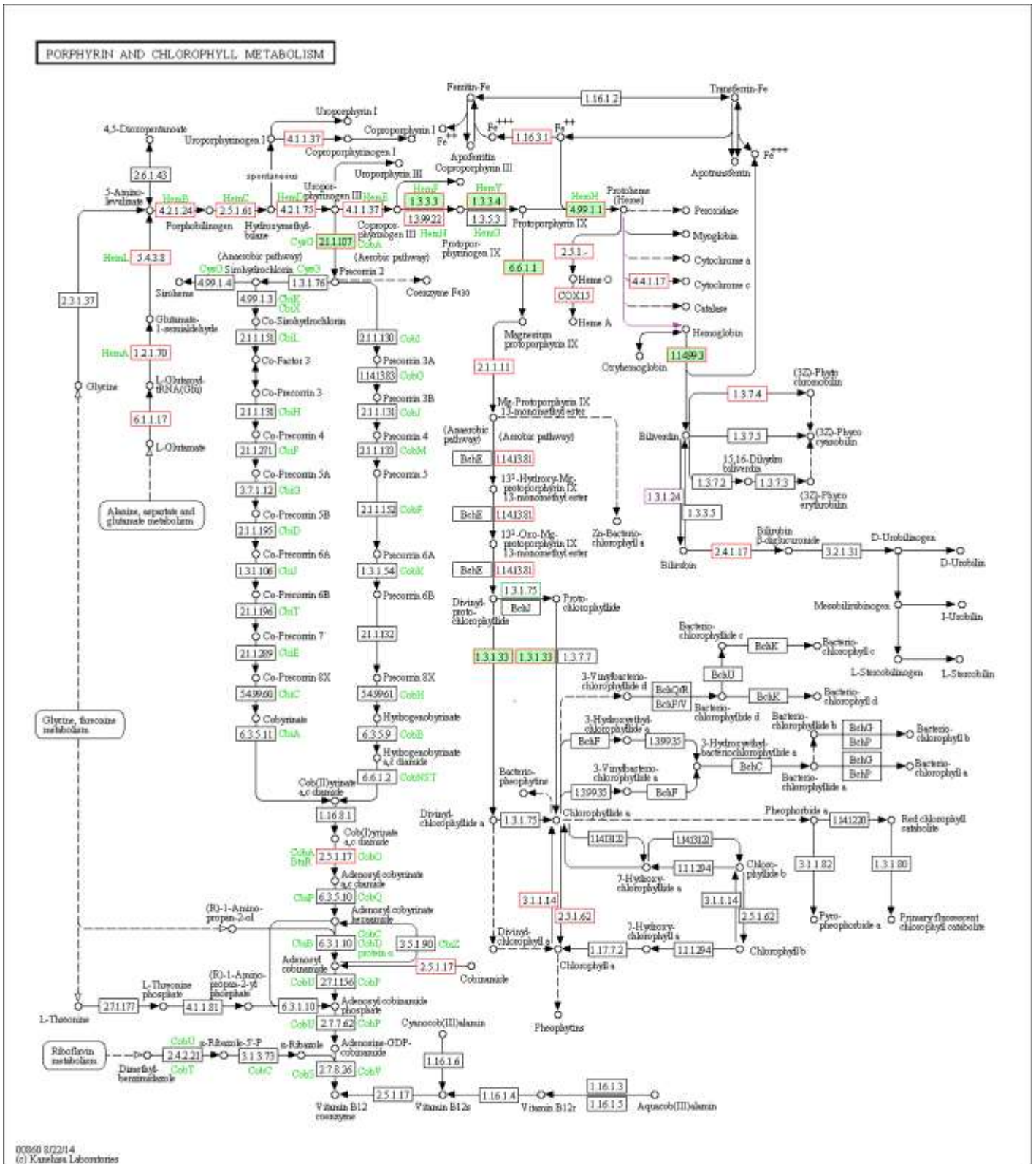
Εικόνα Π.Π.15 Βιοσύνθεση ακόρεστων λιπαρών οξέων



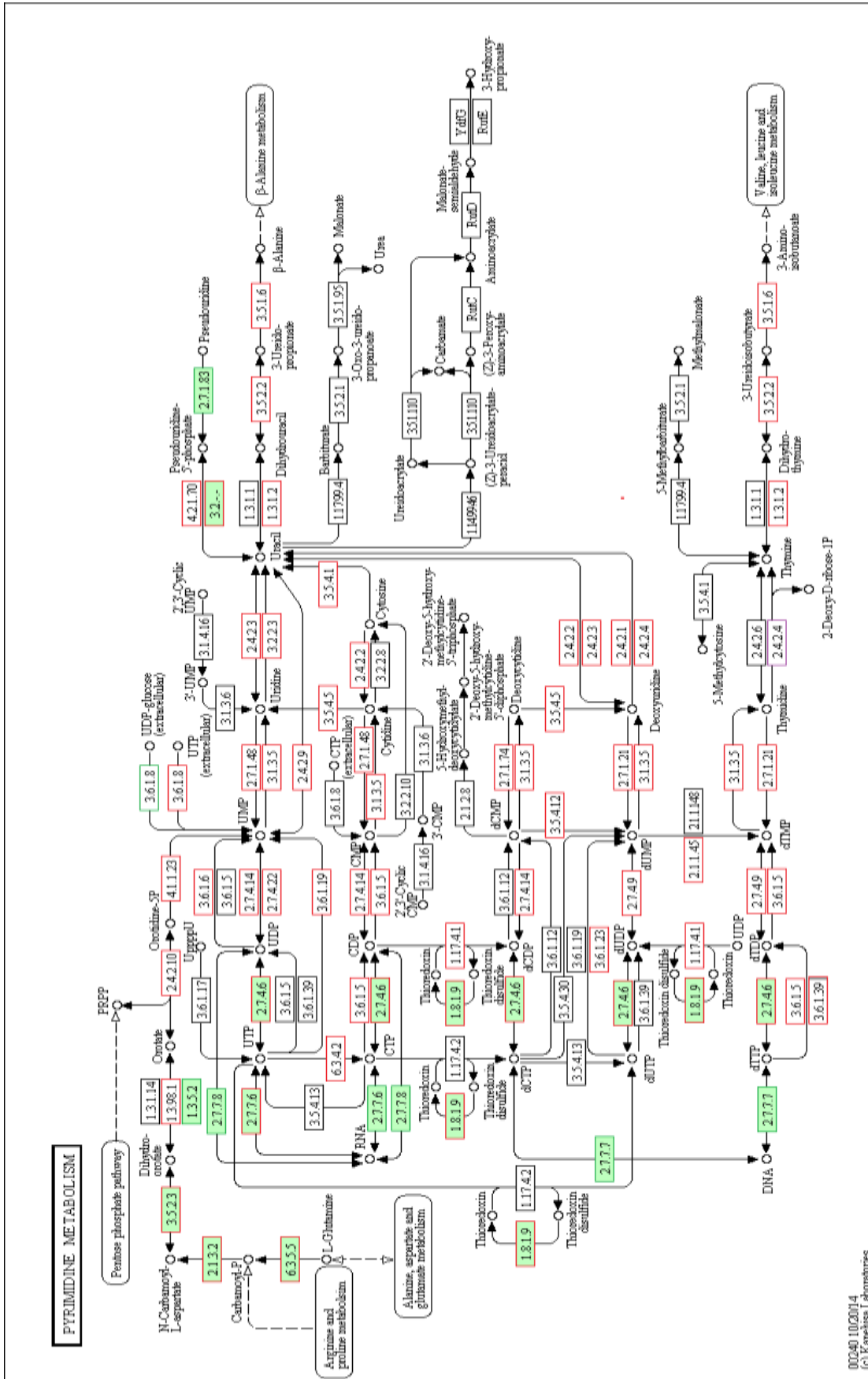
Εικόνα Π.Π.16 Μεταβολισμός βιοτίνης



Εικόνα Π.Π.18 Μεταβολισμός ριβοφλαβίνης



Εικόνα Π.Π.19 Μεταβολισμός πορφυρινών και χλωροφυλλών



Εικόνα Π.Π.22 Μεταβολισμός πυριμιδίνης